Labour supply in urban areas of Bolivia and the role of the informal sector

Pradhan, Menno Prasad

Publication date:
1994

Document Version
Publisher's PDF, also known as Version of record

Labour supply in urban areas of Bolivia

and the role of the informal sector

Menno Pradhan

Labour supply in urban areas of Bolivia

and the role of the informal sector

Menno Pradhan

Labour supply in urban areas of Bolivia
and the role of the informal sector

proefschrift

ter verkrijging van de graad van doctor aan de Katholieke Universiteit Brabant, op gezag van de rector magnificus, prof. dr. L.F.W. de Klerk, in het openbaar te verdedigen ten overstaan van een door het college van dekanen aangewezen commissie in de aula van de Universiteit op vrijdag 14 oktober 1994 te 14.15 uur door

Menno Prasad Pradhan

geboren te Lincoln (USA)

Promotoren:	prof. dr J. Hartog
	prof. dr ir A. Kapteyn
Co-promotor:	dr A.H.O. van Soest

Preface

This thesis marks the completion of my PhD study at Tilburg University. The topic for the thesis arised during the two years that I worked at the World Bank, before joining Tilburg University. The World Bank kindly allowed me to use the data of the Bolivian 1989 household survey for this research. At Tilburg University, I first worked for the Economic Institute Tilburg (EIT). The EIT generously allowed me 3 days a week to work on my PhD. The Netherlands Organization for Scientific Research (NWO) granted me with 2 more years to complete the work. During all this time, Tilburg University provided a stimulating academic environment and interesting colleagues. Chapters 4,5 and 6 of this thesis appeared earlier in the form of CentER discussion papers (no 9311, 9366 and 9429). An article based on chapter 4 has been accepted for publication in Labour Economics. Some individuals in particular contributed to my academic education. They are (in chronological order): my parents, Bernard van Praag, Jacques van der Gaag, John Newman, Arie Kapteyn, Joop Hartog, Arthur van Soest, Rob Alessie, Sanjeev Goyal, Bertrand Melenberg, Gary Fields, Thierry Magnac and Jeffrey James. Without the help of these institutions and individuals this thesis would never have been completed. Thank you.

1. Introduction

Urban labour markets in developing countries are often analyzed within the dual labour market framework. In this framework, two labour market segments are distinguished: (1) a "formal", modern sector, consisting of large scale industries, the public sector and other regulated professions. The wage in this sector is above the market clearing level. This situation is maintained by non-competitive wage setting procedures such as collective labour agreements, minimum wage legislation, efficiency wages, etc. (2) On the other hand the "informal" or traditional sector comprises a wide range of small scale activities. Production generally uses simple technology and requires little capital or formal skills. No barriers to entry exist and earnings are at competitive levels. Rural-urban migration, attracted by high formal sector wages but mostly absorbed by the informal sector, is believed to be the driving force in the expansion of the informal sector in the seventies. A dualistic urban labour market exists if the formal sector maintains its superiority in spite of the excess job seekers.

The question of what exactly constitutes the informal sector is an ongoing debate. The 1972 ILO (International Labour Office) Employment mission to Kenya defined economic informality as a "way of doing things characterized by: (1) ease of entry, (2) reliance on indigenous resources, (3) family ownership of resources, (4) small scale of operation, (5) labour intensive and adapted technology, (6) skill acquired outside the formal school system, and (7) unregulated and competitive markets". Since then different authors have used different definitions. Recurring themes are: small scale and unregulated markets. The first has the advantage that it is easy to measure. The second that it is related closely to economic theory. The two themes often go together: avoidance of regulation is only possible by staying invisible to those enforcing it. For small enterprises the costs of enforcement outweigh the benefits.

Economic policy designed to promote economic growth and/or reduce inequality has to take into consideration the prevailing labour allocation mechanism. Given that labour is the most abundant asset of the poor, it is expected that their principal mechanism to achieve consumption smoothing is through labour supply responses. Restrictions in mobility of labour hinder an efficient allocation of labour over economic activities. The welfare consequences of, for instance, a reduction in public sector wages depend upon the combined effects resulting from adjustments in the supply and demand side of the labour market.

The objective of this study is to increase our understanding of the functioning of the urban labour markets in developing countries, in particular the supply side. We will do so by focusing on one country: Bolivia. Chapter 2 provides background information regarding Bolivia. Bolivia is the poorest country on the Latin American continent. The informal sector provides more than half of total urban employment,

more than in any other Latin country (based on ILO studies in 1985, cf Thomas 1992, p.68). After a period of turbulent economic and political crisis in the first half of the eighties, the economy was stabilized in 1985 and since then has been growing at a very moderate pace. As a direct result of the adjustment program the informal sector expanded strongly in 1985 providing 61 percent of total urban employment in 1986. Since then its size has decreased slightly. We provide an institutional and economic discussion of the urban labour market. Moreover, we discuss the tax system and conclude that income taxes do not play a major role. For the empirical chapters 4-6 we use cross section data from the national household survey in 1989. More information about this survey is given in the last section of chapter 2.

In Chapter 3 we survey the literature of theoretical and empirical studies regarding urban labour markets in developing countries. The theoretical models are all extensions of the, mathematically equivalent, Harris-Todaro (1970) migration model or the Mincer (1976) minimum wage model. Both assume segmented markets, where one segment is covered by minimum wage regulation. Employment in the covered (formal) sector is limited. Unemployment or on-the-job search arises because individuals queue up for the high paying formal sector jobs. The models emphasize that (1) informal sector workers prefer a formal sector job if available and (2) the formal sector offers a higher wage than the informal sector for similar jobs. The empirical studies cast doubt on these assumptions. A significant proportion of informal sector workers does not seem to prefer formal sector employment. A comparison of wage offers leads to mixed results.

Chapters 4,5 and 6 contain an empirical analysis based on the 1989 Bolivian household survey. Chapter 4 contains a comparison of wage offers in the formal and informal sector. We find that, generally speaking, expected wages are higher in the formal sector for males, while they are higher in the informal sector for females. The latter is in contradiction with the simple theoretical models. This suggests that other factors besides the wage offer influence sector participation. The results are influenced by the way in which we corrected for selectivity bias. Selectivity bias arises when workers that have entered into a sector differ in characteristics unobservable to the researcher from those who have not. Formal sector workers, for example, may be risk averse and would therefore not be good entrepreneurs in the informal sector. In this instance, ignoring the selectivity bias would lead to an overestimate of their potential informal sector earnings. We employ two different models, each implying a different selectivity correction. The first one explicitly models the informal sector as a buffer sector, in between non-participation and the formal sector. The second model treats the two sectors identically. Using both models, we find that informal sector earnings are more dispersed than formal sector earnings. Returns to education are higher in the formal sector. Correcting for selectivity bias has a significant effect on the results. Especially the estimate for the constant term in the wage equation is sensitive to the

model specification. On the basis of specification tests we have a slight preference for the first model for males and the second model for females.

Chapter 5 models the sector participation decision in more detail. The results for females in chapter 4 suggest that a simple model where wage offers and rationing of formal sector jobs determine sector participation is not sufficient. We develop a model that distinguishes three factors: wage offers, sector specific non-monetary returns and rationing. Preferences for sectors are not only determined by potential earnings in each sector, but also depend upon non-monetary returns attached to participation in a sector. For example, health insurance in the formal sector or additional freedom in the informal sector may be a consideration in the sector participation decision. Rationing arises from labour market segmentation. It is modelled as the probability that an individual is not able to enter the formal sector. Information on search, including on the job search, is used to identify the model. It is assumed that search is an indication of rationing. The model is estimated using simulated maximum likelihood. We find that males attach higher non-monetary returns to participation in the formal sector. For females, on average, non-monetary returns are not significantly different across sectors. The probability of rationing varies around 30 percent and decreases with education. A simulated increase in the unemployment rate by 5 percent had hardly any effect on the preference of working for males. For females, however, the preference of non-participation increased, suggesting that the discouraged worker effect plays a role.

The model presented in chapter 5 predicts a decrease in the female participation rate if economic conditions worsen. Macro-economic figures, however, suggest the opposite. One possible explanation is the presence of intra-household effects. Females may decide to start working because the earnings of their husband have fallen. In chapter 6 we analyze the joint labour supply decision of two adults households. We follow a structural approach, in which we specify a quadratic household utility function. As in the previous chapter we allow for participation in two sectors and for non-monetary returns that influence the sector participation. We find that a change in the males' own or the females' wage offer affects the labour supply of males very little. Female labour supply is negatively affected by a rise in the wage offer for males, with a cross elasticity of -0.3. A simulation exercise, in which we decreased the formal sector wage offer by 10 percent, resulted in an average decrease of household consumption by 5.0 percent and an expansion of the informal sector by 5.9 percent. One should realize that such an exercise only considers labour supply responses. Taking into account demand factors might lead to a greater decrease in household consumption and a smaller increase of the informal sector, because the decrease in the demand for informal sector services would be taken into account.

Conclusions are in chapter 7. In this chapter, we try to reconcile the empirical results from chapter 4 to 6 an see whether the results combined yield useful information for policy analysis.

2. The Country of Focus: Bolivia

2.1 Introduction

This chapter provides an overview of the history and the economic situation of Bolivia[1]. Moreover, the dataset used in the empirical chapters will be discussed. Bolivia is an appropriate country to study within the dual labour context since the urban informal sector is large, providing about half of total employment. Bolivia is a small, but very diverse country. It is landlocked, bordered on the north and east by Brazil, to the south by Paraguay and Argentina, to the southwest by Chile and to the west by Peru. It has an area of 1,098,581 square kilometres and is divided into nine provinces. In 1989 the country had 7.1 million inhabitants[2]. The country is diverse in its climate and vegetation. The lowlands, found in the east of the country, record an average temperature of 25 °C. Stanta Cruz is the main city in this region. The valleys are found in the middle of the country and vary in altitude between 1000 and 3000 meters. The average temperature in this part of the country is 15 °C. The principal cities in this region are Chocabamba and Sucre (the capital). The highlands (altiplano) are located in the western part of the country and have an average altitude of 3500 meters. The area comprises the cities of Oruro, Potosi and La Paz (seat of the government). About 52 percent of the population lives in this area. Average temperatures range from -3.5 °C in June to 24.5 °C during November and December.

In the last decades, economic performance has been poor and Bolivia is now the poorest countries on the Latin American continent. GNP per capita in 1989 equalled US$ 620. This is substantially lower than in the neighbouring countries. A similar story holds for social indicators. The illiteracy rate among adults was 26 percent in 1985 and life expectancy at birth was only 54 years in 1989. All neighbouring countries score much better on these indicators. The important sectors in the economy of Bolivia are the primary sector and the service sector. The primary sector accounted for 36 percent of GDP in 1989 (of which 22 percent constitutes agriculture and 14 percent mining). The share of the service sector in GDP was 47 percent in the same year. Hydrocarbons and mineral exports accounted for 70 percent of total exports in 1987.

---

[1]This chapter is partly based on findings that were made during a visit to Bolivia in February 1993. Special thanks are due to those who helped me to acquire this information. They are: Guillermo R. Alborta V. (Labour ministry), Juan Carlos Aguilar (World Bank), Silvia E. de Pabón (CEDLA), Susana Arze O., Javier Comboni (both UDAPE) and Manuel Contreras and his staff at UDAPSO. I am responsible for any errors.

[2] Figures reported in this study do not include the latest results of the 1992 census. The 1992 census reported a population of 6.4 million.

The organization of the chapter is as follows: section 2.2 provides a short overview of the history of Bolivia until 1985. Bolivia's economy deteriorated rapidly in the first half of the eighties. In 1985, a successful stabilization package was implemented. The groundwork for the current economic situation was laid in 1985 and therefore, to analyze the current economic situation, the period starting from 1985 should be taken into consideration. Section 2.3 discusses the stabilization program in broad terms. Many of the laws influencing the labour market were instituted along with the stabilization program. Sections 2.4 and 2.5 discuss the institutional framework. The former gives an overview of the tax structure that was implemented while the latter gives an overview of the regulation affecting the labour market more specifically. The data that will be used in the empirical chapters of this study are discussed in section 2.6. Section 2.7 concludes.

2.2 History

Early History. Around A.D. 600 the region around lake Titicaca was dominated by the Aymará and Inca civilization. The Aymarás dominated the region south of the lake, the Inca's the northern side of the lake. By the 10th century much of the Aymará civilization had been conquered by the Incas. Spanish colonization started in the 16th century. The discovery of the mineral wealth, such as the vast silver and mercury deposits of Potosi, made the central Altiplano the administrative seat of the Spanish colonial rule. In the 17th century Potosi was the wealthiest and most populated city with around 160,000 inhabitants in 1650, more than London and Paris together at that time. The first uprising against the Spanish rule started in 1809 with a declaration of independence in La Paz. The rebellion failed, but the resistance against the Spanish rule remained.

In 1825, after a long war, the region known as Alto Perú was granted independence. A few years later, the country changed its name into its present one, Republic of Bolivia, after one of its liberators: Simón Bolívar. Sucre, close to Potosi but with a milder climate, was named capital city. The post-independence period (1825-1880) can be characterized by the violence and corruptness of the successive dictators. A positive exception was the 10 year rule of Andrés de Santa Cruz (1829-1839) who tried to reinstitutionalize the ancient Incan political order and unified Peru and Bolivia by force. His ambitions were halted by Chile which feared the political and commercial rivalry of a powerful Peruvian-Bolivian confederation. The subsequent rulers interchanged office frequently and as a result the political power of Bolivia in the region diminished. In 1879 Bolivia became engaged in a conflict with Chile which resulted in the loss of its coastline.

By the early 1900s, tin mining started to play a predominant role in the economy, taking over the lead from silver. In this period the political party consolidation took place and the country enjoyed an unprecedented period of economic growth and political stability. The economy remained largely confined to mining and related

activities. The vast majority of the population (70 percent) continued to live in a feudal rural economy and was excluded from political decision making (until 1952 less than 3 percent of the population were entitled to vote). The political power remained in the hands of the mining elite, represented in both the liberal and the conservative party.

In 1932 the Chaco war with Paraguay broke out and Bolivia lost a large part of its territory. The costs of the defeat were substantial. Perhaps as many as 60,000 Bolivians died in the conflict and the foreign debt reached 500 billion pounds. The embarrassing defeat triggered the forces leading to the 1952 revolution. The traditional social system of Bolivia had been undermined by the large mobilization of men to the front and the faith in the civilian politicians had eroded. In 1936, the military seized power. The military leaders, favouring a dominant role of the state in the economy, nationalized Standard Oil of Bolivia and named it "Yacimientos Petrolíferos Fiscales Bolivianos" (YPFB).

In the years leading to the revolution, governments interchanged power frequently. The military intervened into politics on a regular basis. The different regimes all favoured a strong role of the state. The differences were in orientation, from military socialism enforcing reforms from above, to conservative representing the interests of the establishment. In this period, the US increased its influence in Bolivia to safeguard its supply of strategic minerals. At the time, Bolivia was the largest tin exporter in the world. Supported during the periods of military socialism, unionization of labour increased, especially within the mining, railway and education sector. New political parties were formed. One of them was the "Movimiento Nacionalista Revolucionario" (MNR) , a party with a nationalistic orientation.

Contemporary history. The populist 1952 revolution elevated the MNR to head Bolivia's first truly nationalistic revolutionary government. Victor Paz Estenssoro was installed as president. The MNR sought broadly based, fundamental reforms in society. Universal suffrage, nationalization of mines and land reforms were among the first acts of the revolutionary government. An educational reform act in 1953 promised education for all. To finance the reforms, Bolivia became increasingly dependent on foreign aid, despite the initial nationalistic orientation of the revolution. In the short run, the assistance guaranteed the economic survival of the revolution. In the long run, however, the conditional US aid forced the government to moderate its initial policies. Overall, the major redistribution of wealth implied by the agrarian reforms initially destabilized the economy. As shown in figure 2.1, between 1952 and 1958, inflation soared and GDP fell at an average annual rate of 2 percent. The economic crisis increased foreign dependency. In 1956, ideological struggles resulted in a breakup of the MNR party.

In the years from 1956 to 1964 the MNR fragmented and the political support for the government weakened. Economic crisis forced the government to install an IMF supported stabilization program in 1956. The unpopular economic plan fragmented the MNR party. Paz, who headed the subsequent governments was the only figure who was acceptable to all parties. Realizing that the democratic support for the revolution was fading, Paz paved the way for a military takeover in 1964.

At first the military rationalized their new rule as a way to protect the revolution. However, in 1968, after the first elections, they decreed one official party and moved to open repression. The rule was conservative and relied on foreign investments and the middle class for support. The policy of "disciplined labour", which was achieved by at times violent confrontation with the miners, destroyed the labour autonomy. In 1969 a socialist military rule took over promoting "revolution from above". Labour unions were allowed to reorganize and the subsidiary of US Gulf oil was nationalized. In 1971 the conservative Banzer took over by military coup. Banzer relied on repressive ruling to restore social order.

Increased foreign aid, favourable international market conditions for minerals and relative stability lead to an average GDP growth of 5 percent over the 1960s and 1970s. Large public investments and job patronage increased the role of the state in the economy (estimated at 33 percent of GDP in 1978). Public investments exceeded private and there was a large private capital outflow. The economic expansion of the 1970s came to a standstill in 1978. Commercial banks started to question Bolivia's debt service capability. By the time it had become evident that the quality of the public investments was often dubious and the oil reserves were not as ample as initially portrayed. Combined with the international debt crisis, Bolivia experienced a severe drop in foreign capital inflow. Economic crisis and US pressure to restore a civilian rule led in 1977 to the downfall of Banzer. After fraudulent elections with military interventions in 1978, 1979 and 1980, Siles assumed office as Bolivia's first civilian president in 14 years.

The lack of external capital flows and the failure to implement a successful stabilization package led to an economic crisis of unprecedented severity in the first half of the 1980s. The inability to borrow money on international markets and the fact that the economy was declining sharply in real terms made it impossible to satisfy the high demands that were made on the new, democratic government. The tax system virtually collapsed, with total government revenues falling from about 9 percent of GNP in 1981 to about 1.3 percent of GNP in the first half of 1985. The huge public sector deficit, approaching 25 percent of GDP by 1985, was increasingly financed through issuing money. Inflation surged, reaching 28,000 percent at an annual rate during the first nine months in 1985. GDP fell 10 percent from 1980 to 1985, 24 percent in per capita terms. The government attempted six stabilization packages. However, these were ad hoc, not well coordinated and opposed by the unions. Due to lack of funds, payments to foreign creditors were

suspended in 1984. The economic crisis forced the Siles government to call for early elections in 1985.

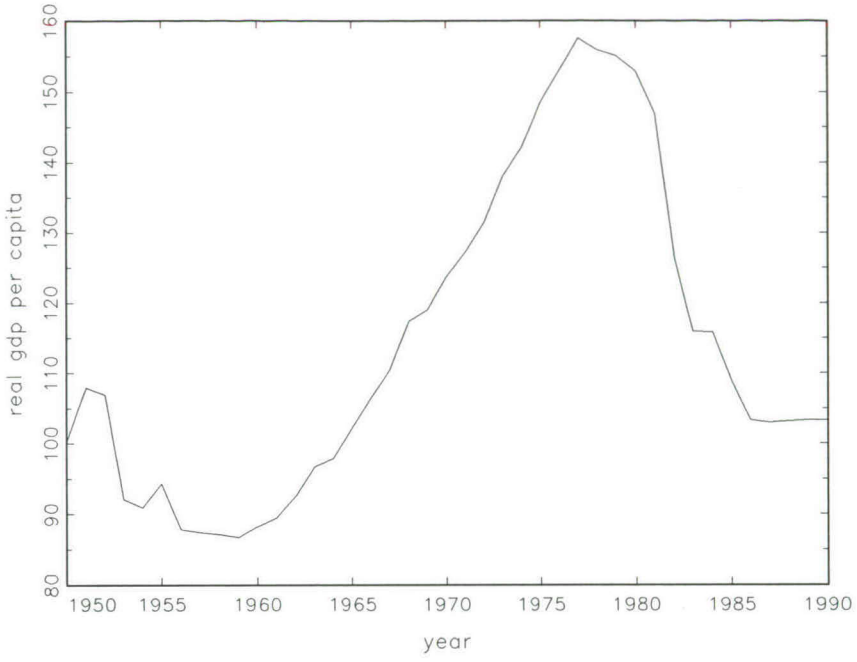

Figure 2.1 Real GDP per Capita from 1950 to 1990. (1950=100 Source: 1950-1985 Summers,R. and A.M. Heston (1988), 1985-1990 World Bank documents)

2.3 The 1985 Stabilization

The successful stabilization program was carried out by the newly elected center-right government of Victor Paz Estenssoro[3]. The agenda of the so called "New Economic Policy" was far reaching and included a massive devaluation of the local currency (93 percent), trade liberalization, fiscal reforms, internal price decontrol and decentralization or privatization of public enterprises. Government expenditures were reduced to levels financeable by available funds. The austerity package had drastic consequences in all sectors of the economy. The sharp rise in public sector prices, especially that of oil, raised the cost of living and increased government revenues. A virtual halt of government investments, a tight freeze on pubic sector wages at depressed levels and a moratorium on foreign debt services decreased the

[3] For a overview of Bolivia's economic crisis and subsequent stabilization program see Morales and Sachs (1988).

expenditure side of the government budget. Approximately 32,000 COMIBOL (the state owned mining company) workers lost their jobs. The Emergency Social Fund, an employment generation project, was created to run parallel with the adjustment program to combat its social costs (Newman, Jorgensen and Pradhan 1991). The adjustment policy was successful in that it stabilized the economy. Within 10 days after the program was announced the inflation was halted and prices actually began to fall. However, the restoration of economic growth took longer. In the first year GDP fell by 2.6 percent. In 1987, growth picked up again and GDP increased by around 2.7 percent, still less than the rise in population, but a dramatic improvement over the first half of the 1980s.

The subsequent years were economically stable with an average GDP[4] growth of 2.8 percent in the period 1987-1990. Even thought this was a reasonable performance in aggregate terms, the rise in per capita terms was insignificant since population grew at a rate of 2.8 percent p.a. In 1991 GDP growth was strongly positive. Overall GDP growth was 4.1% which resulted in a per capita growth of 2.1 percent. The main sectors contributing to this growth were mining (7.6 percent), agriculture (7.2 percent) and manufacturing (6.7 percent). The economy remained reasonable stable in the period following the adjustment: inflation averaged around 17 percent over the period from 1987 to 1991. The deficit of the non-financial public sector averaged around 2 percent of GDP over the same period. The current account balance deficit remained around 9 percent of GNP.

2.4 Tax Structure

The extremely low tax revenues before the stabilization program induced the Paz Estenssoro government to introduce a completely new system of taxes. Before 1985 government revenues had been largely dependent on exports by state enterprises and income taxes. Declining world prices for minerals and an overvalued exchange rate discouraged exports and drove production into illegal channels. Hyperinflation made it almost impossible to collect income taxes. As taxes were collected in local currency it was difficult to correctly index taxes and even a small delay in payment made the revenue virtually worthless. The new system, which was designed in 1986 and implemented in May 1987, relies heavily on consumption taxes to generate government revenues. The new tax system has been responsible for an impressive rise in government revenues. Tax revenues increased from 2.9 percent of GDP in 1985 to 6.2 percent in 1987.

---

[4]Reported GDP figures exclude the value added of coca production. Although difficult to quantify, various studies show a declining role of coca in the economy. According to the estimates, the contribution of coca decreased from 26% of GDP in 1986 to 6% in 1991. The decline is mostly contributed to a fall in world prices.

The main sources of government revenues are (1) taxes and transfers from the state petroleum company (YPFB), (2) value added taxes and (3) custom duties. Revenues from YPFB are the largest source of revenue for the government although the relative share has been declining slowly. Revenues through YPFB dropped from 62 percent of total revenues in 1986 to 43 percent in 1989. The share of the revenues from value added tax (VAT) has been rising steadily. In 1989 VAT payments equalled 23 percent of total revenues. Since the introduction of the new system, custom duties have varied around 11 percent of total revenues. In the following paragraphs we will discuss in more detail the taxes that directly affect the operation of the urban labour market.

There is a value added tax on all goods and services, excluding real estate, export activities and interest payments. This tax was implemented in 1987 at a rate of 10 percent. In March 1992 the rate was raised to 13 percent. The "Specific Consumption Tax" (ICE) levies an additional tax on certain luxury goods. These goods include alcoholic beverages, tobacco, cosmetics and perfumes. The tax rate varies between 30 and 50 percent. For electricity there is an additional tax of 10 percent. Furthermore there is a "Transaction Tax" (IT). The rate was 1 percent until 1989 and 2 percent thereafter. The difference with the VAT is that taxes on intermediate transactions cannot be deducted.

Not all firms include value added tax in their sales. Small firms that qualify for the "Simplified Tax Regime" are exempted from VAT payments. In lieu, they pay a small fixed amount based on their estimated annual income. Businesses qualify for the simplified tax regime if they have an annual working capital of Bs 9600 (in 1989 1Bs=0.37US$) or less, have annual sales of Bs 48,000 or less or sell goods and services with a unit value of Bs 100 or less. Most of the firms that operate in the informal sector fall under this regime. These firms do not issue receipts for their sales. Revenues from this tax are relatively small, in 1989 this tax accounted for 0.4 percent of total treasury revenues.

The system provides an incentive for firms to remain small and not to register for VAT payments. Small businesses can operate more competitively as they do not have to pay VAT. The government promotes compliance with the tax rules by imposing sanctions on those businesses failing to pay the appropriate taxes. For example, in 1989, 3705 businesses were closed for failing to pay taxes. Furthermore, the government sponsored publicity campaigns appealing to national patriotism. In general, the presumption is that large businesses will be noticed by tax law enforcement officers. For small businesses, the administrative costs of collection do not offset the revenues.

The system also provides an incentive for larger businesses to register for VAT. The system works as follows: only VAT paying businesses are entitled to issue receipts on their sales. Individuals can use these receipts to deduct their VAT

payments on purchases from their income tax. Income is taxed at the same rate as the VAT. In principle, if an individual does not save and spends all income at VAT paying firms, he or she does not pay any income tax. As a result, income tax can be seen as a tax on savings and purchases with non-VAT paying firms. A tax credit can be obtained if the tax credits exceed the liabilities. Through this system, VAT paying businesses are favoured by costumers and there is an incentive for firms to register for VAT payments. The system has some deficiencies. A secondary market for receipts has arisen. Individuals can overstate their VAT payments by purchasing additional receipts. Also, the administrative burden to cross-check the receipts issued by the vendors with the receipts declared by the tax payer is great. It is likely that falsified receipts are in use.

The system described above explains the name that is used for income tax: Complementary Tax to the VAT. The tax covers income from all sources, including dividends and interests. The rate is identical to the VAT rate, 10 percent from 1987 to 1992, in 1992 raised to 13 percent. Contributions to social security (5 percent) and to the national housing fund (1 percent) are deductible. In addition, everyone has a personal allowance. Every individual receives a tax credit of 10 percent of two minimum salaries and the tax credit for the VAT payments. For the rest there are few tax deductions. In the empirical chapters we do not make a distinction between before and after tax income. Considering that the Complementary VAT raised only 3.3 percent of total revenues in 1989, this should not cause any problems.

Firms have to pay a "presumed profit tax" which is equal to 2.5 percent of their net worth, the difference between assets and liabilities. In 1990 this tax was raised to 3 percent. Revenues from this tax have been relatively low, in 1989 the contribution of this tax to total revenues was 2.5 percent. The tax induces companies to undervalue their properties for tax purposes. Furthermore, the system favours the service sector over the manufacturing sector, as firms operating in the service sector will usually be less capital intensive.

2.6 Labour Markets

Ever since 1980 the informal sector has accounted for more than half of urban employment. The number of individuals employed in the informal sector is strongly influenced by economic policy. Figure 2.2 presents the share of the urban labour force employed in the formal and informal sector from 1980 to 1989. (UDAPE 1991) During the period of economic crisis in the beginning of the eighties, the share of the informal sector increased only slightly. Job patronage and strong unions kept the share of formal sector employment about constant. There was a sharp rise in informal sector employment just after the reforms: from 1985 to 1986, urban formal sector employment fell by 62,209 jobs, i.e. 14 percent. In the same year the informal sector grew with 116,704 people (UDAPE 1991). This

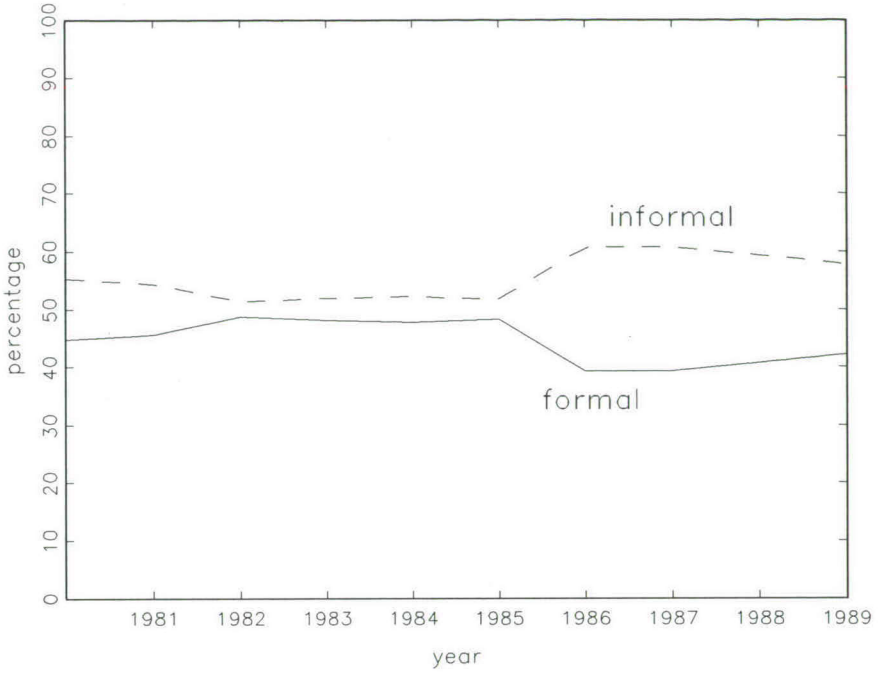

Figure 2.2 Share of formal/informal sector in urban employment (c.f. UDAPE 1991, sector definition not stated).

supports the view of the informal sector as a buffer sector. Income generated from the informal sector is needed in a period of economic recession[5].

There is no one to one relationship between formality and sector of economic activity. The formal and informal sector operate in overlapping sectors of the economy. For males, in 1989[6], formal sector employment was concentrated in the public sector (32 percent), manufacturing (14 percent), transport (13 percent) and construction (12 percent). For females, formal employment is largely concentrated in the public sector (63 percent). For males, informal sector employment is concentrated in commerce (24 percent), personal services (18 percent), construction (17 percent) and manufacturing (14 percent). For females, informal sector employment is concentrated in commerce (69 percent) and personal services (14

[5] Horton (1994) provides an excellent overview of the adjustment process in the labour market in the eighties.

[6]Source: household survey data 1989, definition formal/informal sector based on worker's status (see section 2.7).

percent). The data show that female employment is concentrated in specific sectors of the economy: the public sector for formal sector employment and commerce for informal sector employment. For males, formal and informal activity overlap, it found particularly in sectors of manufacturing, construction and transport.

Most of the labour market regulation affects the formal sector. The basis of the regulation was put forth in the "Ley General de Trabajo" of 1942 (see Rodriguez et al. 1992). In the beginning of the 1980s additional provisions were made, mostly favouring the rights of the workers. In 1985, as part of the New Economic Policy, a number of labour regulations were eliminated. The policy was aimed at reenforcing market powers in the labour market, establishing a closer link between work effort and pay. The policy change (1) allowed for wage negotiations to take place at firm level instead of sector level, (2) eased the hiring and firing of employees and (3) abolished most bonuses (such as subsidized food stores) for public sector workers. The minimum wage legislation was maintained. In the following we will briefly discuss the current regulation facing firms and workers in the formal sector.

The minimum monthly wage was Bs 60 from March 1988 to December 1990. In 1990 it increased to Bs 120. Using the exchange rate of 1989, Bs 60 equalled about US$20. This is fairly low and does not seem to be a major obstacle for entry into the formal sector. In the private sector the enforcement of this law is weak. The minimum wage is mostly offered in the public sector for unskilled workers. For comparison, the average monthly wage in the private sector in 1989 was around 500 bolivianos while in the public sector it was about 155 Bolivianos (UDAPE 1991).

Collective bargaining takes place in the sectors in which the government is involved. Union power in the private sector is small, negotiations take place at firm level and the unions have little impact. This is a direct result of the loosening of the firing rules. Union power used to be strong in the mining sector, but with the malaise in this sector the influence of the unions has diminished accordingly. At this moment, unions are strong in the public sector and in the hydrocarbon sector (YPFB). In these sectors collective labour agreements exist.

In the case of a labour dispute, workers have the right to go on strike. However, for a strike to be considered legal several juridical steps have to be followed. This is a lengthy process that takes at least 30 days. After workers and employers have been in an official state of dispute for at least 10 days, the government will offer to mediate. Within 15 days the government mediator will propose a settlement. This proposal is definitive. If the employers do not accept the proposal they have the right to close the factory and fire all workers. If the workers do not accept the proposal they have the right to go on strike. They should give at least 5 days notice before commencing. In the case of a legal strike the employer has to continue

paying the salary to the worker. Today, there are relatively few strikes as unions have become more tolerant.

There are no unemployment benefits as such. In lieu, employers in some cases have to compensate fired workers. The regulation is as follows: if the worker has been employed for less than three months or if there is a clear understanding at the initiation of employment that the relation will terminate at a certain date, no redundancy pay is required. If employment has lasted for more than three months there is a distinction between dismissal with and without justifiable cause. In case dismissal is with justifiable cause (as defined by the ministry of labour) no compensation is required. If the dismissal is without justifiable cause, the employer has to compensate the worker the equivalent of three monthly salaries plus one monthly salary for every year of service. In practice it is very hard to prove that dismissal is with justifiable cause. Employers can evade indemnity payments by contracting workers for less than three months on a continuous basis.

It is compulsory for all formal sector firms and workers to join social security. It covers health insurance, disability insurance and old age pension. Employers contribute about 15 percent of the wage bill to social security. Workers contribute around 5 percent. All workers have to join the state run National Social Security Fund. Moreover, a sector specific fund has to be joined. The efficiency of the current system has been questioned. Administrative costs use up a large percentage of the flow of funds (estimated at up to 25 percent). Those who can afford it, purchase additional insurance in the private sector. Currently the pension scheme is organized as a pay-as-you-go system. The height of the pension is largely determined by the total contributions at the time and, as a consequence, there is little relationship between any individual's contribution and the benefit. Presently the government is thinking of changing the system into a system of personal capitalization, establishing a closer link between contributions and benefits.

2.7 Data

The empirical research in the following chapters draws upon data from the second round of the 1989 Bolivian household survey (Encuesta Integrada de Hogares, EIH 89), which is one round of a pooled cross section survey[7] conducted by the Bolivian Central Bureau of Statistics (Instituto Nacional de Estadística) since 1976. The survey covers 7246 households in 8 urban centers comprising the three major geographical zones. It is a multipurpose survey. The survey collects individual level information on a wide variety of topics including education, labour supply, health and migration. Household level information is collected on expenditures, non-labour income and housing. Household survey data, in contrast to firm level data, are particularly appropriate for measuring activity in the informal sector since

[7]Every year different households are interviewed.

they are drawn from the entire urban population. Firm level data often do not include non-listed firms (micro-enterprises), of which the bulk of the informal sector consists. The survey collects a measure for household consumption and, for every family member separately, detailed information on labour supply, earnings, education, health, fertility and migration. The labour section of the survey is extensive. It provides information on occupation, earnings, hours worked and search behaviour.

Table 2.1 presents the distribution of the sample over the 8 urban centers and some regional summary statistics. Cochamamba, Santa Cruz and Trinidad have relatively high mean per capita expenditures. The average family size is 6.6. The variables "econ act" and "unemployment" characterize the local labour market. The "econ act" variable denotes the number of people (males and females) working or looking for work in the (random) sample. It relates to the size of the local labour market and may pick up scale effects. It varies between 695 for Potosi to 2486 for St. Cruz. The lowest unemployment rate[8] (males and females jointly) is found in Trinidad (0.053) and the highest in Oruro (0.107). High average per capita expenditures seem to coincide with more favourable economic conditions and low regional unemployment rates.

Table 2.1: Distribution of sample over urban areas plus regional summary statistics

	No of households	monthly p.c. expend.	unempl. rate	econ act
Sucre	583	91.3	0.069	947
La Paz	1338	91.7	0.097	2328
Cochabamba	1192	135.6	0.076	2287
Oruro	776	61.5	0.107	1191
Potosi	439	62.9	0.098	695
Tarija	526	85.1	0.093	979
Santa Cruz	1414	147.1	0.076	2486
Trinidad	996	134.1	0.053	1874
Total	7264	110.1	0.082	

In the subsequent analysis we will use a measure of household net dissavings, defined as total household consumption minus total household labour income. This

[8] Defined as the number of individuals who did not work and searched for during the past week divided by the economically active population (those who searched or worked during the past week).

measure is constructed on the basis of household consumption and individual income data that are collected in the survey. Questions on household expenditures are answered by the head of the household. Total expenditures is constructed by aggregating 27 different categories. These categories constitute a fairly complete set of the main expenditure items. Recall periods vary according to the type of expenditure. A rent is imputed for house owners. Individual earnings are more difficult to measure[9]. Earnings are collected using only one question for each activity. For those involved in household production or in a family business it is difficult to measure earnings. For those participating in the labour market one can make a more accurate estimate of their earnings. Looking at the primary activity only, 97.6 percent of the working prime age males and 85.6 percent of the working prime age females work in an activity for which it is possible to measure earnings.

Further examination of the quality of the data on household consumption, earnings and non-labour income indicates that the calculated household savings measure is a meaningful estimate of actual savings. The mean of monthly household expenditures in the survey equals 682 Bs whereas the mean household income (earnings plus non-labour income) equals 657 Bs. This implies an average household dissavings of 25Bs (0.33 standard error). The figures on household expenditure and income are comparable in magnitude. Savings may be underestimated due to the difficulty of measuring earnings for household production or family labour. Besides, the different time units in which earnings and consumption are measured give rise to measurement error. Regression analysis taking household expenditures as the dependent variable and a constant term, household income and family size as explanatory variables gives the following result (standard errors in parentheses):

$$\underset{\qquad\quad (18.4)}{\text{Expen} = 262.6} + \underset{(0.009)}{0.47\ \text{Inc}} + \underset{(2.61)}{16.5\ \text{FS}} \qquad R^2=0.30$$

There is a significant positive relation between expenditures and income. As income increases, the share of expenditure decreases and that of savings increases.

As noted in chapter 1, the exact distinction between formal/informal sector is an issue of ongoing debate. The information collected in the survey allows for different definitions of the formal and informal sector. To investigate the effect of the use of different definitions we have produced table 2.2. For the working population aged 19 to 65, this table presents a cross tabulation using two commonly used definitions. The first definition uses the size of the enterprise as the primary

[9]The same question about earnings is used for wage workers and self-employed workers. The respondent can choose whether to report by day, week or month. Questions on the average number of hours worked per day and the average days worked per week make it possible to impute hourly earnings.

indicator of formality. If the size of the enterprise is less than 6, the work is classified as informal. However, as is common, independent professionals, such as lawyers and doctors, are classified with the formal sector workers. Household workers and family workers are left unclassified. The second definition is based on the worker's status and corresponds to the definition used by Magnac (1991): wage workers and independent professionals are classified as formal, and self-employed[10] workers as informal. Others, that is employers, and home and family workers, are left unclassified. Of all workers, 4 percent are employers, and are therefore not classified according to Magnac's definition. Both definitions do not classify home and family workers. For this group it is very hard to measure their labour supply and earnings. The change of definition does not affect the sector of classification for 80 percent of the workers (if we include non-classified). Thus, it does not seem to make a large difference which definition is used. In order to make our study comparable with Magnac's we follow his definition in this and the next two chapters. In chapter 5, we will add a capital asset criterion to the sector assignment of self-employed workers.

Table 2.2 Comparison of two definitions of formality
(percentage of working individuals)

	Definition according to size			
Definition according to worker's status	formal	informal	not classified	total
formal	34	16	0	50
informal	0	39	0	39
not classified	2	2	7	11
total	36	57	7	100

The estimation results in chapters 4 and 5 are based on subsamples of 6349 out of 7937 males, and 7293 out of 9028 females. From the original sample, 855 males and 864 females were excluded because they were not working as a result of health problems or because, at the time of the survey, they attended full-time education. 455 males and 648 females were excluded since, according to our definition of formal and informal sector, they could not be classified. These are home or family workers and employers. 19 males and 9 females were excluded because of missing or implausible information on earnings or hours worked. We excluded 10 males

[10]According to the survey an individual is engaged in self-employment if: The person worked without dependence on a boss, managing his own economic unit, with or without the help of family workers or unpaid apprentices, but without using more than two salaried workers.

and 5 females as the education level was not reported, 9 males and 8 females as reported household expenditures were extremely different from reported household income, 226 working males and 175 working females as they did not report earnings, and 14 males and 19 females as of missing information on other household income.

Table 2.3 Sample means by labour market state (sample standard deviations of other variables than dummies in parentheses, see appendix I for a description of variables)

	male		not	female		not
	formal	inform	working	formal	inform	working
education level:						
basic	0.21	0.35	0.24	0.09	0.39	0.29
inter	0.14	0.19	0.13	0.08	0.16	0.15
medio	0.29	0.30	0.33	0.22	0.20	0.29
middle technical	0.04	0.03	0.04	0.09	0.03	0.05
higher technical	0.03	0.02	0.03	0.05	0.01	0.01
normal (teacher)	0.06	0.01	0.02	0.27	0.02	0.03
university	0.19	0.07	0.16	0.17	0.03	0.05
other	0.04	0.03	0.05	0.03	0.16	0.13
married (dummy)	0.79	0.85	0.59	0.55	0.71	0.80
ethnic	0.30	0.39	0.33	0.19	0.46	0.34
age (in years)	35.9	39.7	38.8	33.8	39.3	36.8
	(10.7)	(11.6)	(15.4)	(9.3)	(11.3)	(12.7)
per cap net dissavings (/100)	0.01	-0.04	0.79	0.24	-0.04	0.37
	(2.20)	(1.88)	(1.71)	(2.79)	(1.65)	(1.99)
hourly earnings	2.39	2.58		1.94	2.01	
(primary activity[11])	(4.4)	(4.5)		(1.8)	(4.6)	
hours worked per week	49.7	52.3		38.5	46.9	
(primary activity)	(16.8)	(19.2)		(16.5)	(24.7)	
number of observations	3605	1863	881	1439	1972	3882

The research focuses on labour supply behaviour of individuals from 19 to 65 years of age. Table 2.3 provides summary statistics for the individuals that were included in the estimations. The table uses the definition of formality based on the worker's status. On average, formal workers have a higher education level than informal workers. This difference is the greatest among females. The percentage of individuals reporting to have completed university training is strikingly high. The

[11] The survey collects limited information on the secondary activity. Secondary activity hours and earnings could not be computed in the same way as for the primary job. The number of people with primary and secondary activity in different sectors is small: 2.5 percent of males and 1 percent of females. We therefore ignore the secondary activity in this study.

variable "ethnic" is a dummy variable obtained from the language question: if the respondent commonly speaks another language than Spanish, the variable is set to one. Ethnic workers are overrepresented in the informal sector. The variable "per cap net dissavings" is defined as family expenditures minus family earnings divided by family size. There is no significant difference in the average net dissavings between workers in different sectors. The reason for using this specific other family income measure will be discussed in chapter 4.

The distribution of the natural logarithm of hourly earnings is given in figure 2.3. Hourly earnings are higher for males than for females. For males, the distribution of log hourly earnings is similar in the formal and informal sector. For both sexes, average informal earnings are higher than average formal earnings. For females, the distribution of log formal sector earnings is skewed to the right. The standard deviation of the informal sector earnings is more than twice that of the formal. Both for males and females, the average number of hours worked per week is higher in the informal sector than in the formal.

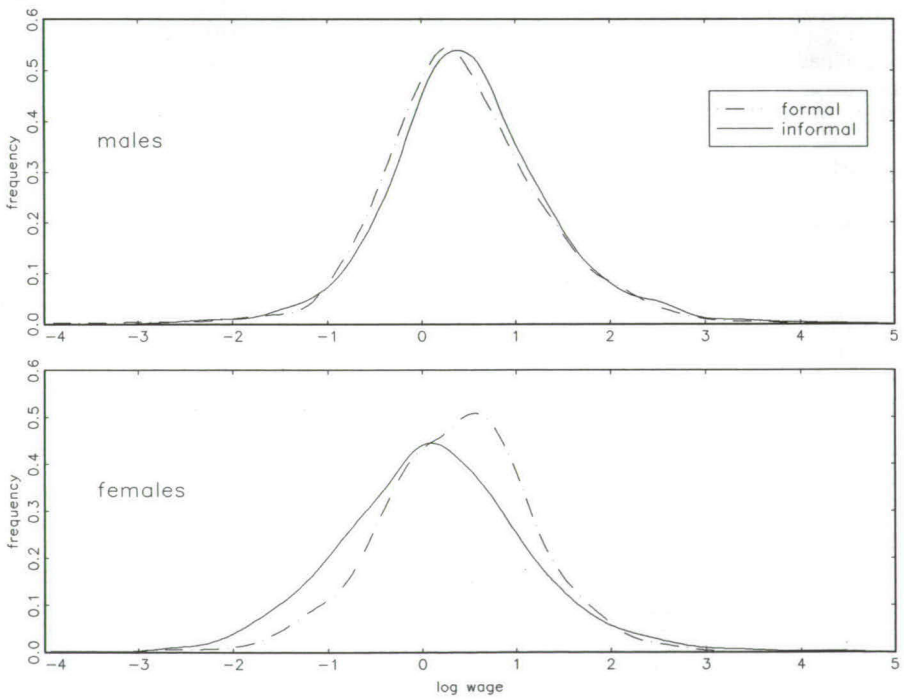

Figure 2.3 Distribution of Log Hourly Earnings[12]

[12] Smoothing is done using kernel density estimation with Gaussian kernel and the normal reference rule bandwidth (see Scott 1992, p. 131).

Due to measurement problems, reported hourly earnings may underestimate the monetary benefits of formal sector jobs compared to informal sector jobs. Fringe benefits are not included in reported formal sector wages. 57 percent of males and 72 percent of females working in the formal sector, reported to have received benefits other than the regular wage. The survey does not collect a monetary equivalent for these. 48 percent of males and 64 percent of females in the formal sector are enroled in social security, the costs of which are partly borne by the employer. Income taxes hardly play a role: in 1989, they represented 3.3 percent of total tax revenues (World Bank 1989). Informal sector earnings may not always be measured net of costs. The questionnaire does not contain enough detail to correct for this. This type of problem may affect estimates of wage differentials between the two sectors, but will not affect the estimates of the (reduced form) sector allocation equations.

To get an idea of the extent to which the data represent national figures, the information contained in the survey can be compared with official estimates from other sources (ILO, 1991). In the survey, 78 percent of all males and 47 percent of all females have a paid job. The activity rates, i.e. the fraction of people working or looking for work, amounts to 84 percent of males and 51 percent of females. Official estimates report activity rates of 93.4 and 26.5 for males and females in the same age group, nationwide in 1990 (i.e. for both rural and urban areas). Official estimates report an average of 44.8 hours of work per week for employees (i.e. a subset of all formal sector workers). The official estimates differ substantially from the ones obtained from the (urban areas only) EIH data. It is not clear to us whether the quality of the official estimates would be "better" than the ones obtained from the EIH survey. At the time of the survey there was no other reliable employment survey in operation. The official estimates are probably also based on the EIH survey, and the difference could be caused by the use of different definitions or by the guestimate that is included in the official estimate for the rural areas.

2.8 Conclusion

Bolivia seems to be an appropriate country to study urban labour markets in developing countries. After the economic crisis in the first half of the eighties, the second half of the eighties has been reasonably stable and one can expect that the labour markets have adjusted to a new equilibrium. In 1989, the informal sector provided employment to about half of the working population in urban areas. But what factors determine the allocation of labour over sectors? In analyzing the urban labour market one cannot neglect the rural labour market and migration. The allocation of labour over sectors depends on how the markets interact. Do labour markets operate competitively or is the allocation a constrained equilibrium? Careful analysis of such issues calls for a theoretical framework. The question of

how to model labour markets for developing countries, both from a theoretical and applied point of view, is the subject of the next chapter.

3. Modelling the Urban Labour Market in Developing Countries

3.1 Introduction

The ideal economic model captures the observed stylized facts, is empirically falsifiable and leads to a better understanding of the functioning of the economy. A model of the urban labour market in a developing country must capture two major empirical facts. (1) First, it must allow for wage dualism, where a formal and informal sector exist simultaneously. This is observed in practically all urban centers in developing countries. The informal sector can take up a sizable portion of the urban labour market. For example, in Latin America, the informal sector labour force ranged from 24 to 63 percent of the total employed labour force in 1985 (Mezzera 1985). Wage dualism arises in a segmented market, where in one sector non-competitive wage setting mechanisms prevail. This could be caused by strong union involvement in one sector, minimum wage legislation, etc. Because entry into the covered (formal) sector is restricted one can observe that workers with identical characteristics are paid differently in different sectors of the economy. In a competitive setting, all workers would move to the sector with the highest wage offer and wage dualism would disappear. (2) Second, the model has to allow for the existence of open unemployment and underemployment[13]. Open unemployment rates are sizable and can run up to double digits. The magnitude of the problem is disguised by the fact that some workers engage in marginal work, working few hours for a very low pay. They are not counted as unemployed, but they are, just like unemployed, rationed because they would prefer a regular job for the going wage to their current position. These workers are also referred to as underemployed.

Continuing rural-urban migration has led to the understanding that the urban and rural labour markets are interrelated. If urban wages increase faster than rural wages this will lead to rural migrants seeking work in the urban labour market. The increased competition that results will affect the expected wage offer for urban residents. The pioneering work in this area is Harris and Todaro (1970) who developed an equilibrium model to explain the persistence of open unemployment in urban areas of a developing country. This model will be discussed in more detail below.

Empirical testing of labour market models requires detailed data on an individual level. The conclusions that follow from theoretical models, such as the Harris-Todaro model, can be confronted with survey data regarding earnings, search for

[13] Open unemployment is defined as population not working, available for employment and actively seeking work divided by total economically active population. For underemployment no standard definition exists. Often an individual is considered underemployed if he or she works involuntarily less than some norm number of hours.

work, migration etc. In testing, one should correct for heterogeneity in the labour force. For example, an observed wage differential can be attributed to differences in skills or restrictions on entry in sectors. While the first is still possible in a competitive setting, the second is only possible when there is labour market segmentation. The scope of the testing depends on data availability. Data that allow to test a complete urban-rural labour market model are rarely available. Usually the tests focus on one specific issue such as wage dualism or urban-rural migration. In this study we will use urban survey data. Our tests will therefore be focused on wage dualism and labour market segmentation. We will abstain from migration issues.

How does a policy intervention in one sector influence participation and earnings in this and other sectors? This is the main question we would like to address with a labour market model. A policy intervention could be anything from a layoff of public sector workers, or the institution of a wage floor, to the promotion of enhanced seeds for farmers. All have a direct effect on the earnings of workers and will thus influence the rest of the labour market. A definite answer cannot be expected since it would require extensive modelling of both demand and supply effects. The focus here will be on the labour supply effects within the urban economy. The objective is to focus on the questions we can answer with the data at hand, hoping that this will ultimately lead to an increased understanding of urban labour markets in developing countries.

The organization of this chapter is as follows: section 3.2 provides a survey of the theoretical models of labour markets in developing countries. A brief overview of the empirical literature is contained in section 3.3. Finally, section 3.4 contains some suggestions on how the theoretical and empirical models can be improved.

3.2 Review of Theoretical Models

We begin with a brief exposition of the influential Harris-Todaro (1970) model. Their objective was to develop a model which could explain the persistence of open unemployment in urban areas of a developing country. The model distinguishes two sectors: (1) an agricultural and (2) a formal urban sector. Agriculture is assumed to be a free entry sector with a given wage W_a. In the formal urban sector a higher wage prevails, W_f. The wage in the formal sector is above market clearing levels and employment in the formal sector is rationed. Let F denote the total amount of urban sector jobs. The labour force, L, is assumed homogeneous. Sector preferences depend on wages only. As the formal sector wage is above the agricultural wage, any individual prefers a formal sector job to an agrarian job. However, not everybody will seek employment in the formal sector as the probability of obtaining formal sector employment diminishes as more individuals queue up for the limited number of positions that are available. Two strategies are

possible: (1) remain in agriculture and (2) seek formal sector employment. In equilibrium, the expected returns for both strategies will equate:

$$E_a(W) = E_f(W) \quad \Leftrightarrow \quad W_a = pW_f \tag{3.1}$$

where E_a is the expectation if the first strategy is followed and E_f for the second strategy. If no formal sector job is found the individual does not work and has zero returns. p denotes the probability of obtaining a formal sector job. This probability equals the total number of formal sector jobs (F) divided by the number of formal sector job seekers, L_s.

$$p = \frac{F}{L_s} \tag{3.2}$$

The labour force is divided into formal sector job seekers plus agricultural workers (L_a).

$$L=L_s+L_a \tag{3.3}$$

Ex post, of the total amount of formal sector job seekers (L_s), a fraction p obtains a formal sector job and a fraction (1-p) remains unemployed (U). Note that in this model on-the-job search from agriculture is not allowed for. The agricultural sector is considered geographically separated from the urban sector. In equilibrium, $p=W_a/W_f$, $L_s=(FW_f)/L_a$ and $U=(1-p)L_s$. As an example, in the case where $W_f=6$, $W_a=3$, $F=200$ and $L=1000$, in equilibrium $L_s=400$, $L_a=600$, $p=1/2$ and $U=200$.

The main contribution of the Harris-Todaro model is that it can explain the allocation of labour over sectors in a developing country as a result of demand side factors. For example, an increase in the urban wage is predicted to lead to an increase in unemployment in the urban sector. Also rural-urban migration can be studied within the context of this model. However, an informal sector is not included in the model. The increased recognition of this sector in the mid seventies led economists to extend the Harris-Todaro model to allow for an informal sector. Such a model would provide a more accurate description of reality. Moreover, if one carries out a simple numerical exercise in terms of the Harris-Todaro Model in which one plugs in estimated values for the exogenous variables, the urban unemployment rate that follows is generally too high (cf Fields 1975 p. 168). The equilibrium unemployment rate equals the ratio of the agricultural and the formal sector wage. According to Fields (1975) observed wage differentials vary from 1/2 to 1/8, yet the highest unemployment rate observed in the seventies in LDCs is 20 percent. A possible explanation for the overprediction could be that some of the formal sector job seekers do not search from unemployment but engage in informal sector work while they continue to search on the job.

Fields (1975) extends the Harris-Todaro model to include an informal sector. Just like agriculture the informal sector is considered to be a free entry sector. Because the informal sector is located in the same geographical area as the formal sector, on the job search is allowed for from the informal sector. In fact, because there are no search costs in this model, all informal sector workers engage in on the job search. Fields (1989) improves upon the 1975 model by (1) making an explicit distinction between ex ante strategies and ex post outcomes and (2) by allowing for search from agriculture. The discussion below presents a simple version of the 1989 model which includes the first feature, but excludes the latter for the sake of clarity. Three strategies exist: (1) remain in agriculture, (2) look for formal sector employment from unemployment and (3) look for formal sector employment from the informal sector. On-the-job search is assumed to be less efficient than full time search from unemployment. Denote the agricultural wage, the formal sector wage and the informal sector wage by W_a, W_f and W_i respectively. The first two are exogenous in the model, the latter endogenous. The allocation of labour in equilibrium is found by equating the expected returns from all strategies

$$E_a(W) = E_f(W) = E_i(W) \quad \Leftrightarrow \quad W_a = pW_f = \theta pW_f + (1-\theta p)W_i \tag{3.4}$$

where θ $(0<\theta<1)$ is the relative efficiency of on the job search relative to search from unemployment. There are several reasons why the efficiency of on the job search from the informal sector may be lower. The limited time that is available for searching or the negative stigma effect associated with working in the informal sector may lower the chance of entry into the formal sector. The number of formal sector job seekers (L_s) equals the labour force queuing up for formal sector jobs weighted by their relative efficiency of search.

$$L_s = L_f^a + \theta L_i^a \tag{3.5}$$

where L_f^a denotes those individuals in the labour force that choose, ex ante, to search from unemployment and L_i^a those who choose to search from the informal sector. As in the Harris-Todaro model the probability p equals the total number of formal sector jobs (F) divided by the number of formal sector job seekers .

$$p = \frac{F}{L_s} \tag{3.6}$$

Of the individuals that choose to search from the informal sector, L_i^a, a fraction $(1-\theta p)$ remains in the informal sector.

$$L_i^p = (1-\theta p)L_i^a \tag{3.7}$$

where L_i^p denotes the labour force, ex post, in the informal sector. The informal sector wage depends on the number of entrants in this sector. The demand for informal sector services, Q, is shared equally among all informal sector workers:

$$W_i = \frac{Q}{L_i^p} \tag{3.8}$$

In this simple version of the model Q is taken as exogenous. In various extensions Q is modelled as a function of the formal and informal sector wage bill. This complicates the expression for a closed form solution and is not necessary for the current argument. Ex ante, the labour force is divided into those who search from unemployment (L_f^a), those who search while working in the informal sector (L_i^a) and agricultural workers (L_a). Ex post, the labour force divides in unemployed (U), formal sector workers, (L_f^p), informal sector workers (L_i^p), and agricultural workers (L_a).

$$L=L_f^a+L_i^a+L_a=U+L_f^p+L_i^p+L_a \tag{3.9}$$

As before, of the workers that choose to search from unemployment (L_f^a), a fraction p obtains a formal sector job and a fraction (1-p) remains unemployed (U). In equilibrium it follows from (3.4) that $p=W_a/W_f$ and $W_i=(W_a-\theta pW_f)/(1-\theta p)$ and from (3.4) that $L_s=(FW_f)/W_a$ and $L_a=((1-\theta)FW_f-\theta Q)/(W_a(1-\theta))$. L_i^a and L_a^p follow directly from (3.5) and (3.7). U and L_a follow from (3.9). As a possible extension of the example introduced earlier one could take $W_f=6$, $W_a=3$, $F=200$, $L=1000$, $\theta=1/2$ and $Q=450$. The equilibrium allocation of labour that follows is $L_f^a=250$, $L_i^a=300$, $L_s=400$, $L_f^p=\mathrm{E}=200$, $L_i^p=225$, $L_a=450$, $U=125$, $W_i=2$ and $p=1/2$. Since a demand for informal sector services is included in the model, in equilibrium, unemployment has decreased as compared to the outcome in the Harris-Todaro model.

The Fields model has some strong implications that can be tested empirically. First, according to the model, all informal sector workers are dissatisfied with their current position and engaged in on-the-job search. This is also referred to as the staging hypothesis, which claims that the informal sector can been seen as a staging post on the way to entering the formal sector. Second, the informal sector is considered to have free entry. All workers can start working in this sector without any major obstacles. Third, in equilibrium, the wage in the informal sector is less than in the formal sector for every individual. Formal sector employment is rationed and more favourable than informal sector work. The next section contains a survey of empirical studies regarding these implications, including results for

Bolivia where available. We will draw in part upon a recent survey of empirical studies by Thomas (1992).

3.3 Review of Empirical Studies

The conclusion from Field's (1975,1989) model that every worker in the informal sector is dissatisfied with his or her current position and seeks work in the formal sector can be easily tested with survey data. Using survey information on "on-the-job" search activity of informal sector workers one can conclude that this proposition does not hold in practice. For example, Welling and Sutcliffe (1984) found low numbers of informal workers actively searching for work on the basis of survey data from South Africa. Strassmann (1987) found that 71 percent of homeworkers in Lima would require a considerable financial incentive to tempt them to move to the formal sector (c.f. Thomas 1992). For Bolivia, in 1989, we find that only 7 percent of the males working in the informal sector are engaged in active on the job search. The rejection of the proposition need not be that strong, however. It is still possible that some informal sector workers would prefer a formal sector job if they got one offered, but do not engage in active search because the potential gain does not outweigh the costs of searching. These, discouraged, searchers are also rationed for formal sector jobs. For Bolivia, if we include discouraged searchers[14] with the active searchers, the search rate for males rises to 29 percent. The high rate of discouragement seems to be justified. Van Lindert (1990), using his own survey data of 1980, finds that migrants change jobs frequently but do not move much towards jobs with a higher status. Chapter 5 contains a more elaborate discussion on job search and discouragement in Bolivia.

Theoretical models consider the informal sector as a free entry, competitively operating sector. In a recent essay, Fields (1990) concludes that "entry into the informal sector is clearly not as easy as the popular image would lead us to believe". Vending spots on the street may be costly, and even places for begging sometimes have to be paid for. Initial capital investments for small firms may be hard to finance as potential entrepreneurs in the informal sector may encounter difficulties in obtaining credit. For Bolivia, a survey among own account workers in 1983 in La Paz showed that 70 percent of the workers financed their initial investment out of personal or family savings (see Sainz et al. 1988). Only 10 percent took a loan and the majority of these were obtained through informal channels. In some sectors of the economy informal workers are organized in associations (gremios). For example, for street vendors, the gremios distribute the vending spots. Although this is not a formal regulation, it may impose barriers to entry. This effect, however, does not seem to be very strong. The same survey

[14]We define discouraged workers as those who answer "There is no work" when they are prompted why they do not look for work. For a more thorough discussion see page 67.

showed that of the informal sector workers in La Paz 40 percent knew about the existence of a gremio for their activity and only 23 percent had actually joined one.

The theory supposes that the wage in the informal sector is lower than in the formal sector for every individual. Empirical studies (see Fields 1990, p 61) have shown that there is a substantial overlap in the income distribution in the two sectors. This does not necessarily imply that the theory is rejected. Skills are distributed heterogeneously across workers. If a highly qualified informal sector worker earns more than a poorly qualified formal sector worker this does not imply that the wage offers in the formal sector are lower. In comparing wage offers one should correct for heterogeneity in productivity.

Not all factors influencing productivity are observed by the researcher. Selectivity bias occurs when there are unobserved factors that influence productivity as well as the selection into sectors. In making predictive statements about wage differentials one has to correct for this bias. As an example, suppose there are two labour market states: working and not working. Selection bias may occur if one predicts the wage of non-participants on the basis of the relationship between wages and individual characteristics of workers only. Suppose, for example, that only individuals with a high level of entrepreneurial skills work. Entrepreneurial skills are unobserved to the researcher. In this case, ignoring selectivity bias would probably lead to an overprediction of the potential wage offer for non-participants. A simple prediction would assume that their entrepreneurial skills are equal to those of workers. To correct for selectivity one has to explicitly model the selection process into sectors.

There is a large literature that deals with empirical testing of wage dualism and labour market segmentation in developing countries. Heckman and Hotz (1986), analyze earnings of prime age males in the labour market of Panama. They test for wage dualism between low income earners and other workers. They recognize explicitly that it is not sufficient to correct for observed heterogeneity in the labour force. They face a problem of sample selection bias because the sample definition for the two subgroups is based on a variable that is endogenous to the wage equations. They correct for this bias and conclude that the earnings functions are significantly different over the two sectors. Gindling (1991), in an analysis of the labour market of San José, Costa Rica, distinguishes between the private formal, the public and informal sector. He only uses data from employed workers, with a positive monthly income. In this case, the sample definition is exogenous but there may be non-random selection into sectors. Because there are three sectors, Gindling employs the multinomial logit model (McFadden 1974) to model the selection process. He finds that there is non-random selection into sectors. Higher educated enter into the public sector while less educated enter into the informal sector. He finds no sector selection bias in the wage equations and proceeds with least squares estimates. The predicted wage differential for an average individual

between the private formal and the informal sector ranges from 15 to 19 percent. The informal sector wage exceeds the formal sector wage for only 9 percent of the workers. These are all workers with less than primary education. Tannen (1991), using data of prime age males from the urban labour market in North-east Brazil, finds that including a Heckman correction term in a (rather restrictive) model eliminates the previously found formal-informal sector wage differential. He concludes that earnings, properly corrected for skills, are not different in the two sectors.

All studies discussed so far use data of working individuals only. Non-participants are left out of the analysis. The analysis is thus conditional on the participation decision. Selection bias problems arise when one wants to make predictive statements for the population as a whole, including non-participants. This problem is not so serious for males. Their participation rate is usually rather high. But for females, ignoring the participation decision could substantially affect the results.

Magnac (1991) develops a model that can be used to analyze the sector participation decision simultaneously with the labour supply decision. His model is an extension of the Roy (1951) model. There are three possible labour market states: not to work, work in the formal sector and work in the informal sector. The individual is assumed to maximize a utility function which depends on consumption and leisure. Under the assumption of a linear budget constraint one can derive an indirect utility function depending on wage and non-labour income which is increasing in the wage. Moreover, one can define a reservation wage under which the individual will decide not to participate. In the empirical model, Magnac models the log of the reservation wage as a linear function of individual characteristics plus an error term

$$\ln(w_R) = Z\beta_R + \epsilon_R \tag{3.10}$$

in which

w_R = the reservation wage,

Z = a vector of individual characteristics determining potential wages, home productivity and taste shifters,

ϵ_R = an error term.

The log wage offers in the formal and informal sector are modelled as a linear function of individual characteristics

$$\ln(w_i) = X\beta_i + \epsilon_i \qquad \begin{matrix} i=1 \textit{ for formal sector and} \\ i=2 \textit{ for informal sector} \end{matrix} \tag{3.11}$$

with

w_i = the wage offer in sector i,

X = a vector of individual characteristics determining potential wages.
ϵ_i = an error term.

Under the hypothesis of competitive markets one would expect individuals to participate if at least one of the wage offers is higher than the reservation wage. If participating, the individual will choose to enter the sector in which the wage offer is the highest. However, if there are restrictions to entry into the formal sector, some individuals with a comparative advantage in the formal sector may decide not to participate in that sector because the costs of entry are too high compared to the advantage of participating. Rationing, in this model, is viewed as costs of entry into the formal sector and is considered in the participation decision. The costs are proportional to wages (waiting queues). The log of costs, c, of entry is given by

$$\ln(c+1) = X\tau + \epsilon_c \tag{3.12}$$

The sector participation decision will now be based on a comparison of the wage offer in the informal sector, $\ln(w_2)$, with the wage offer in the formal sector discounted for the costs of search, $\ln(w_1/(c+1))$. The test of labour market segmentation is based on a comparison of this model with the competitive model in which the costs of search, c, are restricted to be equal to zero ($\tau=0$). Magnac, using data of married woman in Columbia, rejects both labour market segmentation and the equality of the wage equations in the formal and informal sector. He finds that the assumption of competitive markets appears to be an accurate description of the labour market. It seems like there is a selection, as modelled in the Roy model, going on. For example, individuals with a higher education, who have a comparative advantage in the formal sector, have a higher probability of participation in the formal sector. This conclusion is different from the ones discussed earlier and supports a separate analysis of female labour supply.

On the basis of the results from the empirical literature one can conclude that the theoretical models need further improvement. The main criticism of the existing theoretical models seem to be: (1) barriers to entry are also found in the informal sector. (2) Direct evidence using information on search, shows that not all informal sector workers are dissatisfied with their position. (3) Tests for labour market segmentation, based on wage differentials between sectors, produce mixed results. It seems that the main problem is that the theoretical models treat the informal sector as homogeneous without recognizing the wide range of activities and individuals that are found in the informal sector.

How can we reconcile the contradicting conclusions from the theory and the empirical evidence? Two mainstreams can be found in the literature: (1) on the basis of field research in Costa Rica, Fields (1990) proposes to distinguish two sectors within the informal sector: an upper and a lower tier informal sector. The upper tier informal sector workers are, given the constrained choices open to them,

voluntarily in the informal sector. They often acquired their skills in previous formal sector employment. Barriers to entry in this sector exist. The lower tier informal sector is a free entry sector. Workers participate in small scale activities, for example, selling lottery tickets. The latter sector corresponds with the informal sector as it was originally introduced in the Fields (1975) model. (2) Many authors have recognized that an observed wage differential between sectors does not necessarily imply labour market segmentation. Individuals are heterogeneous and may have sector specific preferences that are incorporated in the sector selection decision. Rosenzweig (1988) questions the conclusions based on wage differentials as follows: "Do they suggest barriers to entry -non competing groups- or do they merely reflect compensatory differentials, rewards for unmeasured skills or compensation for unmeasured differences in the disutility of the workplace?". Sector specific preferences could explain why some individuals do not search for formal sector employment even though the formal sector wage offer is higher.

3.4 Suggestions for Improvement

This section list some suggestions for improvement of the existing theoretical and empirical models. The main criticism to the theoretical models seems to be that the informal sector is modelled as fully homogeneous. In the (Fields 1989) model the total demand for informal sector services is shared equally among informal sector workers. One possible way of introducing heterogeneity in this model is the following: assume that individuals differ in their entrepreneurial ability, denoted by ρ (see Brock and Evans 1986, p. 51-53). Entrepreneurial skills are distributed over the population according to the distribution function $H(\rho)$, with support $[\rho_{min},\rho_{max}]$. In agriculture and in the formal sector there are no returns to entrepreneurial ability. In the informal sector incomes rise with entrepreneurial skills according to

$$W_i(\rho) = W_i g(\rho) \tag{3.13}$$

with g a non-decreasing function denoting the returns to entrepreneurial skills. In the previous model g was constant for all ρ. In the present, extended model the expected returns in the informal sector differ across individuals. In equilibrium, the expected returns of the marginal individual who chooses the "search from informal sector" strategy equals the expected returns for the individuals who choose the other strategies:

$$W_a = pW_f = \theta p W_f + (1-\theta p) W_i(\bar{\rho}) \tag{3.14}$$

Here $\bar{\rho}$ denotes the entrepreneurial skills of the "marginal individual". If g is strictly increasing, individuals with entrepreneurial skills greater than $\bar{\rho}$ have the greatest expected returns from the "search from informal sector" strategy. As a result the model predicts a sorting among individuals: individuals with high entrepreneurial skills choose the "search from informal sector" strategy, the others

choose either of the remaining strategies. However, it is not guaranteed that all individuals who choose "search from informal" actually decide to search on the job. It is possible that for some individuals with high entrepreneurial skills, informal sector earnings exceed the formal sector wage ($W_i(\rho) > W_f$). For these individuals there is no reason to search for formal sector work. Thus, a new strategy emerges, namely, to enter into the informal sector without search on the job. W_i remains the variable that ensures equilibrium in expected returns. In equilibrium, the demand for informal sector services equals the total informal sector earnings:

$$Q = L\, W_i(\rho^*) \left\{ \int_{\rho^*}^{\rho_{max}} g(\rho)\, dH(\rho) + (1-\theta p) \int_{\bar{\rho}}^{\rho^*} g(\rho)\, dH(\rho) \right\} \tag{3.15}$$

with $W_i(\rho^*) = W_f$. If g is strictly increasing one can show that compared to the previous model: (1) the informal sector (L_i^p) decreases because Q is shared unequally among individuals. (2) The number of individuals that choose to search from unemployment (L_f^a) increases. This happens because the number of formal sector job seekers is constant ($L_f^a + \theta L_i^{as} = FW_f/W_a$, with L_i^{as} those who, ex ante, choose "search from informal"). Hence a decrease in L_i^a causes L_f^a to increase. (3) The number of individuals that remains in agriculture (L_a) increases. This model has the advantage over the previous one that it allows for some individuals in the informal sector to be better off than in the formal sector. Moreover, the effects of a change in the distribution of entrepreneurial skills in the informal sector can be studied within the context of this model. Other possible extensions, such as an explicit modelling of risk, are reviewed by De Wit (1991, chapter 2).

In the next chapters we improve upon the empirical modelling. We start by an analysis of wage offers in the formal and informal sector. Since we will examine both male and female employment we model the participation decision as well. We experiment with different types of selection models. Next, in chapter 5, we explicitly deal with the critique that observed differences in wage offers do not necessarily imply barriers to entry but may also be indicating compensating wage differentials. Next, in chapter 6, we have a closer look at the participation decision and allow for intra-household effects. All these extensions improve upon analysis of the labour supply decision. Demand side effects are modelled rather parsimoniously. We exploit regional differences in labour markets and individual search information which we interpret as an indication of rationing. More rigorous modelling of the demand side would be worthwhile. This would require a more general equilibrium model approach. However, the available cross section allow for a partial approach only.

4. Formal and Informal Sector Employment

4.1 Introduction

In this chapter we analyze earnings and labour market participation in urban areas in Bolivia, using household survey data from 1989 described in section 2.7. We analyze the determinants of potential earnings in the formal and informal sector, considering, for example, returns to education and the effects of local labour market conditions. Are the wage determination mechanisms in the two sectors different from each other? Moreover, we analyze participation. What factors determine whether an individual works in the formal or informal sector? Are the two sectors competitive and is the difference between potential earnings sufficient to explain the sector someone is in? Or do other factors matter, such as (non-monetary) sector preferences or barriers to entry?

The objective is to find out which labour market model describes the Bolivian urban labour market best. Do the markets operate competitively and is the distinction between formal and informal sector just semantics? Or are the two markets segmented with barriers to entry in the formal sector? Or does the "weakly competitive equilibrium" as defined in Magnac (1991) provide the best description? In the latter, different wage determination mechanisms in sectors exist but there are no barriers to entry. In the analysis we allow for heterogeneity in the labour force and correct for non-random selection into sectors. The analysis is reduced form because the factors (wage offers, preferences and rationing) driving the sector participation decision are not modelled separately. We examine predicted wage differentials between sectors. In the case of competitive markets and homogeneous preferences every individual would move to the sector with the highest returns. In that case, predicted wage differentials would be sufficient to explain the sector participation decision.

Three labour market states are distinguished: not working, working in the informal sector, and working in the formal sector. In a two sector framework, models analyzing labour market segmentation and returns to human capital have been used extensively. See, e.g., Hartog and Oosterbeek (1993) for a developed country and Van der Gaag and Vijverberg (1988) for a developing country. These studies condition on participation in the labour market. Particularly for females, conditioning on participation limits the validity of the results. For three labour market states no established approach exists.

In this study two models based on different approaches towards the informal sector are used. In the first one the informal sector is viewed as an intermediary sector between not working and the formal sector. This stems from the traditional staging hypothesis in, for example, the pioneering work of Fields (1975): formal sector employment is rationed and all informal sector workers would be better off in a

formal sector job (cf section 3.2). In the second approach, competitive markets are assumed and the two sectors are treated symmetrically. Examples are Magnac (1991), who tests for labour market segmentation in Columbia, and Gindling (1991), who analyzes urban labour markets in Costa Rica (cf section 3.3). We consider both approaches and analyze the sensitivity of the results for the choice of model. In particular, estimated wage equations may differ, because the models lead to different corrections for selection bias (Heckman and Hotz 1986). We choose between models on the basis of specification tests.

The organization of the chapter is as follows. In section 4.2, we introduce the two models. Results are discussed and evaluated in section 4.3. Conclusions are mentioned in section 4.4.

4.2 Models

The main objective of this chapter is to analyze factors driving participation, sector choice and earnings in the formal and informal sector. We use two models. The models consist of two wage equations, one for each sector, and two reduced form equations explaining the selection mechanism. They differ with respect to the latter. The number of hours worked is not considered. Wages are hourly wage rates, obtained by dividing total earnings by the number of hours worked.

In the first model, the choice between working in the formal sector, working in the informal sector, and non-participation, is modelled using ordered probit. In the second model, multinomial logit is used. The selection equations are in reduced form, in the sense that the wage rate is not included as an explanatory variable. Wage effects are thus indirectly reflected through, for example, age and education effects. The model is also reduced form in the sense that we consider someone's actual state only. Information on preferred labour market state or job search is not taken into account. We do not disentangle effects through preferences from those through rationing, costs of search, etc.

Ordered probit selection model In the first model, the three labour market states are ordered: participation in the formal sector - participation in the informal sector - non-participation. An underlying latent variable can be interpreted as an indicator of formality. Non-participation includes being engaged in household production which is associated with the lowest level of formality.

We do not explicitly specify an underlying structural economic model leading to this ordering. The economic interpretation is the staging hypothesis of the Fields (1975) model: informal sector employment is inferior to formal sector employment; informal sector earnings exceed unemployment income. In equilibrium, a move from unemployment to the informal sector involves an increase in earnings at the cost of a reduced efficiency of search. In this view, economic formality is

associated with increasing economic activity and a decreasing search effort (see also Todaro 1989, p. 268). In a dynamic setting, new labour market participants or immigrants from rural areas would first accept a job in the informal sector, and simultaneously look for formal sector work.
The formal representation of the model is:

$$Y = Z\delta_1 + \epsilon_3 \qquad \begin{array}{rl} \textit{working in formal sector if} & Y < \alpha_1 \\ \textit{working in informal sector if} & \alpha_1 < Y < \alpha_2 \\ \textit{non-participant if} & \alpha_2 < Y \end{array} \qquad (4.1)$$

The subscript indicating the individual is suppressed. Y is a latent variable, the inverse of the "degree of formality". $\epsilon_3 \sim N(0,1)$, independent of Z. Z is a vector of individual, family, and regional characteristics. Since the system is reduced form (wage rates are eliminated), Z contains all variables in the wage equations. In addition, Z contains taste shifters that do not result from potential earnings differences. By means of normalization, Z contains no constant term.

In the standard version of the model stated above, α_1 and α_2 are constant across the sample. Identification requires only one of them to be constant, however. Keeping the other α constant limits model flexibility: choice probabilities depend on Z through the one index $Z\delta_1$ only; the probability of informal sector employment would depend on Z only through the non-linearity of the distribution function of ϵ_3. This is a general drawback of the standard ordered probit model. To allow for more flexibility, we parametrize α_2:

$$\alpha_2(Z) = \alpha_1 + \exp(Z\delta_2) \qquad (4.2)$$

According to (4.1) and (4.2), the probability of formal sector employment is determined by $Z\delta_1$ only, but the choice between informal employment and non-participation depends on both $Z\delta_1$ and $Z\delta_2$. Incorporating the exponential term in $\alpha_2(Z)$ guarantees $\alpha_2 > \alpha_1$.

Multinomial logit selection model In the second model, no a priori ordering among the three states is assumed. The model can be interpreted in terms of utility maximization. Let Y_i be the indirect utility associated with participation in sector i. We assume

$$Y_i = Z\delta_i + \eta_i \qquad \begin{array}{rl} \textit{working in formal sector if} & \max\{Y_1,Y_2,Y_3\} = Y_1 \\ \textit{working in informal sector if} & \max\{Y_1,Y_2,Y_3\} = Y_2 \\ \textit{non-participant if} & \max\{Y_1,Y_2,Y_3\} = Y_3 \end{array} \qquad (4.3)$$

Here $\eta_i \sim$ EV(I) (an extreme value type I distribution), and η_1,η_2,η_3 independent. Z is the same as in the ordered probit model. Alternative i is chosen if its utility

exceeds that of all other alternatives. Of course, since the actual and preferred state do not necessarily coincide, the interpretation of utility maximization should not be taken literally. If someone prefers but cannot find a formal sector job, Y_1 will be small. Thus Y_i reflects both rationing and preferences.

Normalization requires one δ_i to be constant (we choose $\delta_3=0$). Define

$$\eta_i^* = \max(Y_j) - \eta_i \qquad (j=1,2,3 \quad j\neq i) \tag{4.4}$$

Domencich and McFadden (1975) show that the probability of being in state i equals

$$P_i = prob(\eta_i^* < Z\delta_i) = F_i(Z\delta_i) = \frac{\exp(Z\delta_i)}{\sum_{j=1,2,3} \exp(Z\delta_j)} \tag{4.5}$$

Here F_i is the distribution function of η_i^*, which also depends on $Z\delta_j$, $j\neq i$.

Wage equations The natural logarithm of the potential hourly wage rate w_i in sector i is modelled as

$$\ln(w_i) = X_i\beta_i + \epsilon_i \qquad i=1(\textit{formal}),\ 2(\textit{informal}) \tag{4.6}$$

X contains explanatory variables: personal characteristics (human capital variables), and variables describing the condition of the labour market by urban area. ϵ_i is a normally distributed error term.

Error structure In the ordered probit case, the three error terms ϵ_1, ϵ_2 and ϵ_3 are assumed to be jointly normally distributed with zero mean and a full covariance matrix. The covariance of ϵ_1 and ϵ_2 is not identified, since we observe one wage at most.

In the multinomial logit case, we follow Lee (1982) (see also Maddala 1983, p. 273). Let

$$\epsilon_{3i}^* = \Phi^{-1}(F_i(\eta_i^*)) = J_i(\eta_i^*) \tag{4.7}$$

Here Φ^{-1} is the inverse of the standard normal distribution function. (4.7) implies that $\epsilon_{3i}^* \sim N(0,1)$. Alternative i is chosen if $\epsilon_{3i}^* < J_i(Z\delta_i)$. We assume that, for $i=1,2$, $(\epsilon_i,\epsilon_{3i}^*)$, is bivariate normal with mean (0,0) and covariance matrix Σ_i, with $\Sigma_i(2,2)=1$. These assumptions make it convenient to estimate the multinomial logit model and wage equations with a two step method or maximum likelihood.

Identification Both models are generalizations of two sector self-selection models, of which the model of Roy (1951) is the seminal example. Identification of this

type of models is discussed in Heckman and Honoré (1990) and Heckman (1990). Results in the latter can straightforwardly be extended to our models.

First, the complete model is identified under the specific distributional assumptions on the errors, even without exclusion restrictions on the regressors in one or more of the equations. This changes if distributional assumptions are relaxed and replaced by, for instance, the semiparametric assumptions of independence of errors from regressors and zero error means or medians. For this case, Heckman (1990, p. 314) proves identification under the assumption that at least one of the regressors in the selection part is excluded from the wage equations. This condition is satisfied in our case: Z contains variables referring to family composition and other family income, not included in X. In practice, whether results depend on distributional assumptions, may thus depend on the explanatory power of the additional regressors in the selection equations.

Moreover, semiparametric estimation of the constant terms in the wage equations is harder than estimating slope coefficients. Semiparametric identification of the constant term in, say, the wage equation of sector 1, relies on the assumption that a non-zero population fraction has a probability close to one of being selected into sector 1. As a consequence, semiparametric estimates of the constant term would be inaccurate if few observations have selection probability close to one. In other words, parametric estimates of constant terms may be particularly sensitive to the choice of distributional assumptions.

Net dissavings To allow for an income effect on labour supply, some measure for full income or income excluding earnings is included in Z. For consistency with a life cycle framework (Blundell and Walker 1986), the measure should be corrected for savings: in a two stage budgeting framework, the within period allocation of leisure and consumption is conditional on full expenditures in the same period, or, equivalently, on net dissavings, defined as consumption expenditures minus earnings. We use net dissavings per capita, assuming that the household's net dissavings are shared equally among family members. In this theoretical framework, net dissavings may be treated as an exogenous variable. However, if unobservable factors affect net savings and labour supply, treating net savings as exogenous leads to biased results. We therefore have instrumented for net savings. Non-labour income is the main instrument. Details are is described in Appendix 4A.

Estimation If Z, including net dissavings, is exogenous, the models can be estimated by full information maximum likelihood, or with a consistent two step estimator. The latter can be used to obtain starting values for maximizing the likelihood. Differences between the two sets of estimates can be used to carry out a Hausman test for model misspecification. Details on both estimators are mentioned in appendix 4B.

Allowing for endogeneity of net dissavings as described in Appendix 4A, does not substantially complicate the estimation. In an auxiliary first OLS step, an equation explaining net dissavings is estimated. OLS residuals are then added to the systematic parts of the selection equations. Non-zero coefficients of these residuals indicate endogeneity. See Appendix 4B. Standard errors of ML have been corrected for the uncertainty in the first step estimates, using the results of Newey (1984).

4.4 Estimation Results

Tables with detailed estimation results of two step and ML estimations, are in table 4C1 and 4C2 in the appendix. The following discussion is based on the, more efficient, ML results[15].

Wage equations In Table 4.1 we present estimates for the wage equations. The estimates of the age pattern and the constant term in the informal sector are sensitive to the choice of selection model. For most other explanatory variables the estimated coefficients in both models are comparable in sign and significance. The hypothesis of equal wage equations is clearly rejected for both models and both sexes using a likelihood ratio test[16].

The coefficients on the age variables are significant (at 5 percent level) only for females in the formal sector. A maximum is reached at about 46 years of age. Females in ethnic minorities are significantly underpaid in both sectors according to both models. For males, the signs are the same, but significance levels are lower.

The variables "economic activity" and "unemployment" describe local labour markets conditions. "Economic activity", which is used as a proxy for the size of the local labour market, is positive and significant in all cases. The local unemployment rate has a negative effect on earnings, and the effect is largest in the informal sector. This suggests that the informal sector is more competitive, and that less favourable labour market conditions have a greater negative impact on earnings in the informal than in the formal sector.

Table 4.1 also contains estimates of the error variances of the wage equations and of error covariances between wage and selection equations. The informal sector

---

[15] The terminology in the sequel ignores the preliminary step instrumenting for per capita net dissavings. The two step estimator thus is actually a three step estimator. Similarly, what we call maximum likelihood is not ML of the full model, but the two step estimator with OLS in the first step, and quasi generalized ML on the rest of the model in the second step.

[16] LR test statistics are for males 232 (probit) and 184 (logit) and for females 48 (probit) and 100 (logit) exceeding critical values of $\chi^2(14)$ for every sensible significance level.

Table 4.1. ML estimates wage equations (standard errors between parentheses, * significant at 5% level)

	ordered probit				multinomial logit			
	males		females		males		females	
	formal	informal	formal	informal	formal	informal	formal	informal
cnst	0.042	0.221	-1.777*	0.273	-0.058	-1.961*	-1.887*	1.874*
	(0.174)	(0.278)	(0.288)	(0.331)	(0.168)	(0.302)	(0.278)	(0.387)
age	0.026*	-0.005	0.081*	0.009	0.031*	0.095*	0.083*	-0.024
	(0.008)	(0.012)	(0.012)	(0.014)	(0.008)	(0.013)	(0.012)	(0.015)
age squared /100	-0.007	0.030*	-0.088*	0.000	-0.013	-0.094*	-0.092*	0.034
	(0.010)	(0.015)	(0.017)	(0.017)	(0.010)	(0.015)	(0.016)	(0.018)
inter	0.082	0.052	0.220*	0.256*	0.082	0.077	0.227*	0.275*
	(0.049)	(0.061)	(0.088)	(0.071)	(0.048)	(0.065)	(0.088)	(0.076)
medio	0.245*	0.284*	0.523*	0.360*	0.248*	0.204*	0.540*	0.529*
	(0.041)	(0.054)	(0.075)	(0.064)	(0.040)	(0.058)	(0.073)	(0.069)
midtech	0.440*	0.377*	0.813*	0.566*	0.441*	0.215	0.853*	1.001*
	(0.078)	(0.136)	(0.104)	(0.121)	(0.076)	(0.137)	(0.100)	(0.131)
hightech	0.582*	0.371*	1.028*	0.813*	0.582*	0.143	1.079*	1.375*
	(0.100)	(0.169)	(0.141)	(0.245)	(0.099)	(0.167)	(0.137)	(0.251)
normal	0.078	-0.419*	0.972*	0.125	0.106	-0.610*	1.044*	1.208*
	(0.083)	(0.166)	(0.121)	(0.175)	(0.080)	(0.166)	(0.112)	(0.161)
university	0.827*	0.600*	1.311*	0.718*	0.854*	0.227*	1.371*	1.407*
	(0.048)	(0.080)	(0.110)	(0.117)	(0.046)	(0.091)	(0.102)	(0.127)
other	0.211*	-0.404*	-0.209*	-0.042	0.211*	-0.537*	-0.210*	-0.042
	(0.067)	(0.106)	(0.098)	(0.074)	(0.064)	(0.118)	(0.098)	(0.080)
yrs incompl educ	0.061*	0.046*	0.088*	-0.002	0.062	0.017	0.094*	0.039
	(0.012)	(0.018)	(0.017)	(0.019)	(0.011)	(0.018)	(0.017)	(0.021)
econ active /10	0.172*	0.136*	0.176*	0.151*	0.169*	0.118*	0.177*	0.117*
	(0.023)	(0.034)	(0.032)	(0.035)	(0.022)	(0.036)	(0.032)	(0.038)
unemployment *10	-0.664*	-0.889*	-0.704*	-0.854*	-0.676*	-1.218*	-0.727*	-0.904*
	(0.090)	(0.133)	(0.121)	(0.149)	(0.087)	(0.145)	(0.121)	(0.157)
ethnic	-0.061	-0.073	-0.183*	-0.145*	-0.066*	-0.028	-0.194*	-0.311*
	(0.034)	(0.045)	(0.055)	(0.048)	(0.033)	(0.048)	(0.054)	(0.053)
σ_i	0.905*	0.986*	0.684*	0.954*	0.878*	1.110*	0.689*	1.138*
	(0.012)	(0.026)	(0.011)	(0.017)	(0.011)	(0.033)	(0.013)	(0.039)
σ_{3i}	0.704*	0.653*	-0.080	0.268*	0.655*	-0.923*	-0.138*	0.803*
	(0.023)	(0.040)	(0.076)	(0.062)	(0.022)	(0.048)	(0.068)	(0.070)

wage variance is larger than that of the formal sector wage. This corresponds to the notion that the formal sector is more regulated, leading to a smaller earnings dispersion. The covariances determine the differences between the potential wage distribution for the whole population, and the actual wage distribution among those selected in the sector concerned. It is the latter which is reflected in the sample, and the former which we try to analyze. For males, the positive coefficients for the formal sector imply negative selectivity effects. This is not what one would expect. For males in the informal sector, the selectivity effects are positive according to both models[17]. This means that, for given observed characteristics, those with highest unobserved informal sector skills indeed work in the informal sector. For females, the multinomial logit model leads to exactly the opposite results: the selection effect is positive for the formal, and negative for the informal sector. The difference between the two models seems substantial here: according to the multinomial logit model, the selectivity effects are much stronger than according to the ordered probit model. The correlation coefficients for the informal sector are surprisingly high, explaining the huge selection effects, particularly in the multinomial logit model.

Education level is incorporated as follows: first, dummies are used to indicate the highest level of courses attended. Seven levels are distinguished. "Basic", which is the lowest education level, is the excluded reference category. Second, for those who did not complete the course, we used the deviation between the level attained and the level if completed, expressed in years. This deviation is zero if the course is finished and negative otherwise. Those who followed training that was not classified, are included in the estimation. For them, the dummy variable OTHER is set to one and "years of incomplete education" to zero.

For a discussion of the effect of education on the wage offers we turn to table 4.2. For each level of (completed) education, we have calculated the predicted log wage. For the formal sector, the wage pattern as a function of education level is robust with respect to the chosen model. The relative increase in wage offer from lowest (basic) to highest (university) type of education is 0.84 for males and 1.34 for females. For the informal sector the estimates vary substantially with the choice of the selection model. In particular, the estimated average wage levels are quite different according to the two models. This may be explained by the problems with semiparametric identification of the constant term (cf. the discussion in section 4.3). Returns to education are lower in the informal sector than in the formal sector, as we would expect under the segmented labour market hypothesis.

[17]Because the probability of non-participation is small for males, this statement is valid for the ordered probit model as well.

Table 4.2. Effect of education on predicted log wages (age=35, city=La Paz, ethnic=0. Standard errors in parentheses)

		Males		Females	
	years educ	ord probit	mult logit	ord probit	mult logit
FORMAL					
basic	5	0.638	0.603	-0.300	-0.390
		(0.048)	(0.047)	(0.144)	(0.130)
inter	8	0.720	0.685	-0.080	-0.163
		(0.052)	(0.050)	(0.137)	(0.127)
medio	12	0.883	0.851	0.223	0.151
		(0.042)	(0.040)	(0.110)	(0.101)
midtech	13	1.078	1.044	0.513	0.464
		(0.078)	(0.076)	(0.100)	(0.095)
hightech	15	1.220	1.186	0.729	0.689
		(0.097)	(0.095)	(0.124)	(0.122)
normal	17	0.717	0.709	0.672	0.655
		(0.079)	(0.076)	(0.059)	(0.058)
univ	20	1.465	1.457	1.012	0.981
		(0.046)	(0.044)	(0.072)	(0.069)
INFORMAL					
basic	5	-0.142	-0.677	0.099	0.828
		(0.067)	(0.083)	(0.069)	(0.099)
inter	8	-0.090	-0.600	0.356	1.104
		(0.073)	(0.091)	(0.081)	(0.107)
medio	12	0.142	-0.473	0.460	1.357
		(0.060)	(0.084)	(0.070)	(0.111)
midtech	13	0.234	-0.462	0.666	1.830
		(0.138)	(0.150)	(0.119)	(0.164)
hightech	15	0.229	-0.534	0.912	2.204
		(0.170)	(0.178)	(0.242)	(0.271)
normal	17	-0.561	-1.287	0.224	2.036
		(0.178)	(0.190)	(0.166)	(0.203)
univ	20	0.458	-0.450	0.817	2.236
		(0.092)	(0.124)	(0.119)	(0.178)

Figure 4.1 shows the estimated population distribution of potential earnings in both sectors according to both models. There is a substantial overlap in the distribution of earnings in informal and formal sector, caused by variation in observed and unobserved characteristics. For males, formal sector wages are on average higher than informal sector wages. Examining the expected wage offer (not accounting for the unobserved characteristics) for all individuals separately, one finds that for all males in the sample, the expected wage in the formal sector exceeds that in the informal sector, according to both models. In this respect, the distribution in the whole population differs from the distributions in the selected subpopulations. In particular, the selection effect in the informal sector is quite strong: only those with relatively high potential informal sector earnings will indeed work in the informal sector. This is why the sample average of earnings of informal sector workers

exceeds that of formal sector earnings. This finding yields some support for the Fields (1975) model: on average, males would be better off in the formal sector.

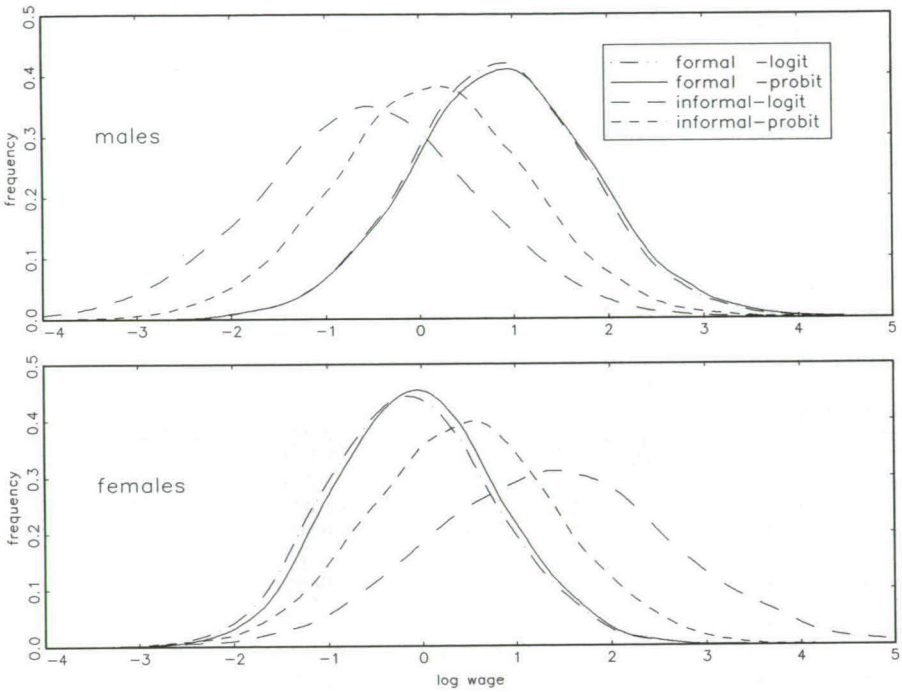

Figure 4.1 Estimated Distribution of Hourly Earnings[18]

For females however, we find exactly the opposite results. According to the multinomial logit model, expected earnings are higher in the informal sector for all females in the sample. Using ordered probit, we find that for 9 percent of the sample the predicted wage offer in the formal sector is higher. These are exclusively females with normal (mainly teachers) or university training. These results indicate that we cannot explain sector participation on the basis of expected wages and restrictions to entry in the formal sector only. Account must be taken of differences in preferences for the two sectors.

Selection equations Interpreting separate parameter estimates in the selection equations is not very useful. Human capital variables appear in various ways (age

[18] Wage offers include random draws of the error terms. Smoothing is done using kernel density estimation with Gaussian kernel and the normal reference rule bandwidth (see Scott 1992, p. 131).

and age squared, for example), and selection probabilities are often determined by more than one linear combination of the regressors. In table 4.3, we therefore present effects of marginal changes of some of the characteristics on the state probabilities, for the average male and female. Signs and significance levels of the effects according to the two models are similar in most cases.

The effect of family composition is captured trough YOUNG, PRIME and OLD. The presence of young children in the household (YOUNG) significantly decreases the probability that the female works in the formal sector, and increases her probability of non-participation. The presence of other prime age individuals (PRIME) implies that the male or female concerned may not be the main earner in the household. This significantly reduces the female's informal employment probability, and increases her probability of non-participation. The latter is also true for males, at the cost of the formal employment probability. A negative impact of PRIME on participation was expected. It is not clear why PRIME affects the choice between formal and informal employment. The presence of family members older than 65 (OLD) is significant in one case only; it increases the male's probability of informal employment.

Individual characteristics are MARRIED, AGE and ETHNIC. For females, being married strongly and significantly reduces employment probabilities in both sectors. For males, being married reduces the probability of non-participation and increases that of formal sector employment. The probability of informal employment increases significantly with age for both males and females. In return, the female's probability of not participating decreases with age, and so does the male's probability of formal employment. This may reflect a cohort effect rather than a pure age effect. Since we included age and its squared, marginal effects at other age levels may be different. Note that age effects are a combination of direct and indirect effects: they include those through the wage rate. For ethnic minorities (ETHNIC=1), the probability of informal employment is significantly larger than for others. For males, this must be a direct effect since ETHNIC hardly affects wages. The relative magnitude of the effect alters with the model chosen.

Education level enters through various variables, and it is not possible to present a single partial derivative. Effects of education on the participation probabilities according to the two models are virtually identical. In general, they are as expected: probabilities of formal and informal employment increase and decrease with education level, respectively. We find a strong positive impact of 'normal' training on the formal employment probability. This type of education includes teachers training college, typically used in the formal sector.

Per capita net dissavings has, for both sexes, a significantly positive effect on the probability of non-participation. Interpreted in labour supply terms, this implies that leisure is a normal good. An increase of net per capita dissavings by 1 percent

Table 4.3 Predicted partial derivatives of probability of participation for an average individual for ordered probit and multinomial logit selection model (standard errors in parentheses; *: significance at 5% level).

ORDERED PROBIT		males			females	
	formal	informal	not working	formal	informal	not working
young	-0.005	0.004	0.001	-0.009*	-0.005	0.014*
	(0.004)	(0.004)	(0.003	(0.003)	(0.003)	(0.004)
prime	-0.012*	-0.004	0.016*	-0.002	-0.015*	0.017*
	(0.003)	(0.004)	(0.003	(0.003)	(0.003)	(0.004)
old	-0.028	0.025	0.003	-0.0075	0.02	007
	(0.015)	(0.016)	(0.013)	(0.012)	(0.013)	(0.017)
married	0.091*	0.019	-0.110*	-0.152*	-0.115*	0.267*
	(0.017)	(0.018)	(0.013)	(0.010)	(0.012)	(0.015)
netdissav pc /1000	-0.624*	-0.189	0.814*	-0.205*	-0.250*	0.455*
	(0.176)	(0.197)	(0.135)	(0.097)	(0.107)	(0.141)
age * 100	-0.478*	0.534*	-0.057	-0.046	0.547*	-0.501*
	(0.070)	(0.073)	(0.057)	(0.051)	(0.054)	(0.068)
econ act /10	-0.007	0.003	0.004	0.004	0.026*	-0.029*
	(0.011)	(0.010)	(0.008)	(0.007)	(0.007)	(0.009)
unemployment *10	-0.005	-0.060	0.065*	-0.144*	-0.007	0.151*
	(0.041)	(0.040)	(0.030)	(0.027)	(0.029)	(0.038)
ethnic	-0.062*	0.049*	0.013	-0.056*	0.062*	-0.006
	(0.014)	(0.013)	(0.010)	(0.011)	(0.009)	(0.013)
MULTINOMIAL LOGIT		males			females	
young	-0.003	0.004	-0.001	-0.010*	-0.004	0.014*
	(0.004)	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)
prime	-0.015*	-0.003	0.018*	-0.004	-0.014*	0.018*
	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)
old	-0.027	0.032*	-0.006	-0.003	-0.001	0.004
	(0.016)	(0.013)	(0.011)	(0.011)	(0.017)	(0.019)
married	0.103*	-0.026	-0.077*	-0.154*	-0.094*	0.248*
	(0.017)	(0.015)	(0.010)	(0.010)	(0.013)	(0.016)
netdissav pc /1000	-0.703*	0.129	0.574*	-0.207*	-0.360*	0.566*
	(0.185)	(0.169)	(0.108)	(0.097)	(0.144)	(0.155)
age *100	-0.585*	0.630*	-0.045	-0.061	0.669*	-0.608*
	(0.073)	(0.069)	(0.046)	(0.052)	(0.067)	(0.072)
econ active /10	-0.002	0.001	0.002	0.007	0.027*	-0.034*
	(0.011)	(0.010)	(0.007)	(0.007)	(0.009)	(0.010)
unemployment *10	0.023	-0.111*	0.088*	-0.134*	0.000	0.134*
	(0.042)	(0.039)	(0.026)	(0.026)	(0.037)	(0.041)
ethnic	-0.065*	0.058*	0.007	-0.054*	0.098*	-0.043*
	(0.014)	(0.013)	(0.009)	(0.011)	(0.012)	(0.014)

increases the probability of non-participation by 0.9 percent for males and by 0.7 percent for females. For females, both formal and informal sector participation probabilities are negatively affected. For males, this only is the case for the formal sector.

"Economic Activity" (EA) and the unemployment rate (UR) characterize the individual's local labour market. A larger labour market significantly increases the probability that a female works in the informal sector, at the cost of non-participation. A high unemployment rate increases non-participation, for both males and females. This may reflect a discouraged worker effect, or, to some extent, an indirect wage effect. For females in particular, results suggest that the discouraged worker effect in the formal sector is substantial.

Specification tests The models are non-nested. In the (extended) ordered probit model and the multinomial logit model, the number of slope parameters is the same. Vuong (1989) developed a test for the null hypothesis that the two models are equally close, in the Kullback-Leibner sense, to the true data generating process, against the alternative that one of the models is closer. This test is quite general, since it does not assume that either of the models represents the truth.[19] For males the null hypothesis cannot be rejected at a 5 percent level. The likelihood of the ordered probit model is larger than that of the multinomial logit model, but the difference is too small to reject equality of expected log likelihoods. For females, the null hypothesis is rejected in favour of the multinomial logit model. For both sexes, multinomial logit outperforms ordered probit in the percentage of correctly predicted observations: multinomial logit predicts the labour market state of 59.0 percent of females and of 62.4 percent of males correctly. For the ordered probit the percentages are 56.1 and 59.2, respectively. The special case of a constant α_2 of the ordered probit model is clearly rejected by a likelihood ratio test.

Model specifications were separately tested using (generalized[20]) Hausman specification tests (Hausman 1978), based upon comparison of (efficient) ML estimates with (consistent but inefficient) two step estimates. Computational details

---

[19] Performing the test is easy here, since we have strictly non-nested models (Vuong 1989, p. 317). We ignored the preliminary step of estimating the dissavings equation. Since estimates of ML standard errors which account for this preliminary step, are virtually identical to standard ML standard errors ignoring the first step uncertainty, we do not think that this leads to biased results.

[20] The tests are generalized in the sense that we have no specific alternative in mind for which ML is inconsistent while the two step estimator remains consistent. This may affect the power, but does not affect the size of the test.

are in appendix 4B[21]. All model specifications are rejected at the 1 percent level. Given the large number of observations and the previous experience in the literature with these kind of models, this is no surprise. For example, the model of Magnac (1991) also fails to pass similar tests. The values of the test statistic lead to the same conclusions as the likelihood values: for males, the ordered probit version of the model performs better, for females the multinomial logit model.

4.5 Conclusions

We have analyzed earnings and labour market participation in urban areas of Bolivia using two different approaches. The first one, the ordered probit model, explicitly models the informal sector as a buffer sector, between non-participation and the formal sector. The second one, the multinomial logit model, does not impose any ordering. We have compared the two models using estimation results, specification tests and comparative statics.

We have generalized the standard version of the ordered probit model, allowing one of the thresholds to vary with characteristics. The model thus gets the same flexibility as a multinomial model without ordering. This generalization appears to be an improvement: the special case of a constant threshold is strongly rejected for both sexes, using e.g. likelihood ratio or Wald tests.

The models are consistent with a life cycle framework in the treatment of income other than earnings. Moreover, we allow for endogeneity of the other income measure, net dissavings. This appears to be an improvement also: LM tests confirm that endogeneity is significant. Estimating the model without allowing for endogeneity yields some quite different results: the impact of unearned income changes sign, and the conclusions about selectivity in the wage equations change substantially. Allowing for endogeneity, we find that leisure is a normal good: The income effect on non-participation is clearly positive. For males, the income effect on formal sector participation is clearly negative, whereas it is insignificant for the informal sector. For females, both income effects of participation are significantly negative. Although the difference between the two models is in the way sector participation is modelled, the estimated probabilities correspond reasonably well, and depend on the regressors in a similar way.

---

[21] Again, we ignored the preliminary OLS estimation of the net dissavings equation. Given the negligible differences between uncorrected and corrected ML-standard errors, we do not think that this is important. Two step method standard errors for the wage equation are corrected for heteroskedasticity and imputing the estimated Mill's ratio following Newey (1984).

Contrary to Gindling (1991) we find that accounting for selectivity substantially affects the wage equation estimates. The direction of the selectivity effects is the same according to both models. In the informal sector, however, the magnitude of the selectivity effect substantially depends on the model which is chosen. As a consequence, the two models lead to substantially different predictions of potential informal sector earnings of non-participants and formal sector employees. Using specification tests for non-nested models, we tend to prefer the multinomial logit model for females. For males, we find slightly more support for the ordered probit model, but the difference is insignificant.

A number of conclusions appear to be robust with respect to the chosen model. The dispersion of potential earnings in the informal sector is much larger than in the formal sector (see fig. 4.1). This is in accordance with the notion that the informal sector is strongly heterogeneous (Fields 1990). Wages in both sectors are higher in larger local labour markets, and lower in areas with high unemployment. In both sectors females of ethnic minorities are generally underpaid, while for males this effect is insignificant. In accordance with economic theory, returns to education in the formal sector exceed those in the informal sector. Predicted wages are higher in the formal sector for males, which supports the view in theoretical models (Fields 1975) that formal sector employment is preferred to informal sector employment, but that formal sector jobs are rationed. For females, however, the opposite holds. This could indicate that it is not sufficient to look at wages only to explain the participation decision of females. It may also indicate relatively more favourable labour market conditions for females in the informal sector. The result is not in line with Magnac (1991) who could not reject weakly competitive labour markets for married woman in Columbia. Further analysis, which disentangles the effects through rationing and differences in preferences, is presented in the next chapter.

In general, it thus seems reasonable to conclude that the impact of most explanatory variables is robust with respect to the specification choice. This is not the case for the exact magnitude of selectivity effects, particularly in the informal sector. Looking at two models instead of just one might then be a first step towards a more robust view on the segmented labour market hypothesis.

Appendix 4A. Endogeneity of Net Dissavings

In the standard life cycle model, net dissavings and the error terms in the labour supply equations (due to future uncertainty only) are uncorrelated. Net savings may, however, be endogenous as a result of unobserved heterogeneity. We add a reduced form equation for net dissavings to account for this:

$$netdissav = Q\ \xi + u \qquad u \sim N(0,\sigma^2) \tag{4A.1}$$

Here netdissav is per capita net dissavings, and Q is a vector of household specific variables.[22] Endogeneity arises if u is correlated with the errors in the selection model.

In the ordered probit model, we assume that u and the error ϵ in the selection equation are bivariate normal. This implies that the conditional expectation of ϵ given u is linear in u. In the multinomial logit model we make a similar assumption, concerning the distribution of the error terms η_i in (4.3), conditional on u:

$$\bar{\eta}_i = \eta_i - \gamma_i u \sim EV(I) \text{ and } \bar{\eta}_1,\bar{\eta}_2,\bar{\eta}_3 \text{ independent} \tag{4A.2}$$

Normalization requires $\gamma_3=0$. If $\gamma=(\gamma_1,\gamma_2)=(0,0)$, the specification is as described in section 4.3, and net savings are exogenous.

Preliminary OLS-estimates of the dissavings equation (4A.1) were used for both models. Results are mentioned in table 4A.1. As expected, the impact of other household income is positive and strongly significant.

---

[22] Q includes household non-labour income. Since this is not in the vector Z of explanatory variables in the selection equations, identification of the model is guaranteed.

Table 4A.1 OLS Results net dissavings equation (variables AGE through "years incompl educ" refer to head of household)

Dependent variable: Household Net dissavings, per capita		
R^2 : 0.03		nr of obs: 7220
	estimate	std err
INTERCEPT	-284.94	136.76
OTHER INCOME	0.26	0.02
LA PAZ	-19.76	48.32
COCHABAMBA	2.07	49.09
ORURO	-130.06	53.17
POTOSI	-81.44	61.24
TARIJA	53.55	59.26
ST CRUZ	-8.42	49.31
TRINIDAD	-82.44	52.61
YOUNG	-1.16	8.64
PRIME FEMALE	-17.15	11.33
PRIME MALE	-38.38	12.33
OLD	-15.55	38.38
AGE	8.49	5.58
AGE SQUARED	-0.03	0.06
MARRIED * AGE	95.80	33.12
ETHNIC	26.22	28.44
INTER	71.98	36.46
MEDIO	134.85	32.78
MIDTECH	227.42	71.56
HIGHTECH	289.16	81.12
NORMAL	102.69	57.63
UNIVERSITY	220.36	41.16
OTRO	-115.89	46.69
YRS INCOMPL EDUC	21.91	9.56

Appendix 4B. Estimation and Testing

Estimation The likelihood function for the full information maximum likelihood estimator is given below. We start with the case in which we do not take account of possible endogeniety of net dissavings. For workers the contribution to the likelihood consists of the marginal density of the wage times the conditional probability of participation in the individual's actual sector. For non-workers, the likelihood contribution equals the marginal probability of non-participation. All marginal and conditional probabilities can easily be computed, only requiring univariate normal distribution functions. In case of the multinomial logit model, this is the motivation for using the transformation given in (4.7).
Log-likelihood ordered probit model:

$$\sum_{formal} \ln\left\{f_1(\epsilon_1)\int_{-\infty}^{\alpha_1} f_{31}(\epsilon_3|\epsilon_1)d\epsilon_3\right\} + \sum_{informal} \ln\left\{f_2(\epsilon_2)\int_{\alpha_1}^{\alpha_2(Z)} f_{32}(\epsilon_3|\epsilon_2)d\epsilon_3\right\} + \sum_{\substack{non\\participant}} \ln\left\{\int_{\alpha_2(Z)}^{\infty} f_3(\epsilon_3)d\epsilon_3\right\} \quad (4B.1)$$

Log-likelihood multinomial logit model:

$$\sum_{formal} \ln\left\{f_1(\epsilon_1)\int_{-\infty}^{J_1(Z\delta_1)} f_{31}(\epsilon_3|\epsilon_1)\,d\epsilon_3\right\} + \sum_{informal} \ln\left\{f_2(\epsilon_2)\int_{-\infty}^{J_2(Z\delta_2)} f_{32}(\epsilon_3|\epsilon_2)\,d\epsilon_3\right\} + \sum_{\substack{non\\participant}} \ln(P_3) \quad (4B.2)$$

Here f_1 and f_2 are the marginal density functions of the error terms in the wage equations. f_{31} and f_{32} are conditional (normal) density functions of ϵ_3, given ϵ_1 and ϵ_2, respectively. J_1, J_2 and P_3 are defined in (4.5) and (4.7).

The two step estimator is computationally the least demanding. The first step consists of estimating the selection submodels (ordered probit or multinomial logit) by maximum likelihood. This yields consistent estimates of δ_1, δ_2 and (in the probit case) α_1. Using these estimates, the wage equations can be estimated by ordinary least squares including correction terms to account for endogenous selection. The wage equations for the formal and informal sector with correction terms are stated below. For the ordered probit specification they are (cf. Idson and Feaster 1990)

$$\ln(w_1) = X_1\beta_1 - \sigma_{13}\frac{\phi(\hat{\alpha}_1 - Z\hat{\delta}_1)}{\Phi(\hat{\alpha}_1 - Z\hat{\delta}_1)} + \epsilon_1^* \qquad \textit{(formal sector)} \quad (4B.3)$$

and

$$\ln(w_2) = X_2\beta_2 + \sigma_{23}\frac{\phi(\hat{\alpha}_1 - Z\hat{\delta}_1) - \phi(\hat{\alpha}_2 - Z\hat{\delta}_1)}{\Phi(\hat{\alpha}_2 - Z\hat{\delta}_1) - \Phi(\hat{\alpha}_1 - Z\hat{\delta}_1)} + \epsilon_2^* \quad \textit{(informal sector)} \quad (4B.4)$$

where $\hat{\alpha}_2 = \exp(Z\hat{\delta}_2)$.
For the multinomial logit specification we get (cf. Lee 1982):

$$\ln(w_i) = X_i\beta_i - \sigma_{i3}\frac{\phi(J_i(Z\hat{\delta}_i)}{F_i(Z\hat{\delta}_i)} + \epsilon_i^* \qquad \begin{matrix} i=1 \textit{ if working in formal sector} \\ i=2 \textit{ if working in informal sector} \end{matrix} \quad (4B.5)$$

Here ϕ and Φ denote the standard normal density and distribution function, respectively. F_i is defined in (4.5). The errors ϵ_1^* and ϵ_2^* now have zero conditional mean, given that the corresponding wage rate is observed. The equations can be estimated by OLS, using only those observations for which the relevant wage rate is observed. Including these correction terms is a straightforward extension of Heckman's (1974) two stage method. Ordinary least squares yields consistent estimates but, unless $\sigma_{i3}=0$, standard errors computed in the usual way are estimated inconsistently, because of the heteroskedasticity of ϵ^* and because the parameters in the correction terms are replaced by their estimates.

Estimation in case of endogenous net savings The more general model allowing for correlation between the error u in (4A.1) and the errors in the selection equations, can be estimated as above, adding one preliminary step: first, the auxiliary equation (4A.1) is estimated with OLS. Then the model itself is estimated conditional on the residual in (4A.1). For the ordered probit selection model, the approach is the same as in Smith and Blundell (1986). The conditional distribution of ϵ_3 is again normal, and its mean is a linear function of u. Inclusion in Z of an unbiased estimator of u, the OLS-residual $\hat{u}$, corrects for the non-zero conditional expectation of ϵ_3. Significance of the estimated coefficient of $\hat{u}$ implies rejection of exogeneity of net savings.[23]

A similar procedure can be followed for the multinomial logit case, based on the linearity assumption implied by the specification given in (4A.2). The OLS-residual from (4A.1) is now added to the regressors in both selection equations in (4.3), and thus replace the unobserved error terms in (4A.2). A joint test for the significance of the two corresponding coefficients indicates whether endogeneity plays a role. In

[23] Using the terminology of Gourieroux and Monfort (1989), this estimator procedure can be interpreted as a quasi-generalized ML-estimator, since the unknown parameters in (8) have been replaced by their estimates.

both models, the ML-standard errors in the second step are inconsistent in case of endogeneity, because the parameters in (4A.1) are replaced by their estimates. Correction is possible following the procedure given by Newey (1984).

The derivation of the Hausman test statistic The Hausman test is based on the statistic $(\hat{\beta}_{ml}-\hat{\beta}_{ts})'[V(\hat{\beta}_{ml}-\hat{\beta}_{ts})]^{-1}(\hat{\beta}_{ml}-\hat{\beta}_{ts})$, with $\hat{\beta}_{ml}$ the maximum likelihood estimator of β(kx1) and $\hat{\beta}_{ts}$ the two step estimator. Let $L(y_i|X_i,\beta)$ be the log likelihood contribution for observation i. The maximum likelihood estimator is defined as

$$\hat{\beta}_{ml} = \underset{\beta}{argmax} \sum_i L(y_i|X_i,\beta) \qquad (4B.6)$$

Let $\beta=(\beta_1,\beta_2)$ and $y=(y_1,y_2)$. If the log likelihood contribution can be split up according to $L(y_i|X_i,\beta)=L_1(y_{1i}|X_i,\beta_1)+L_2(y_{2i}|X_i,\beta_1,\beta_2)$ the two step estimator is defined as

$$\hat{\beta}_{1ts} = \underset{\beta_1}{argmax}\sum_i L_1(y_{1i}|X_i,\beta_1) \quad , \quad \hat{\beta}_{2ts} = \underset{\beta_2}{argmax} \sum_i L_2(y_{2i}|X_i,\hat{\beta}_{1ts},\beta_2) \qquad (4B.7)$$

The calculation of $\hat{V}(\hat{\beta}_{ml}-\hat{\beta}_{ts})$ is not trivial. Although $\hat{V}(\hat{\beta}_{ml})-\hat{V}(\hat{\beta}_{ts})$ is a consistent estimator of the asymptotic variance covariance matrix of $\sqrt{T}(\hat{\beta}_{ml}-\hat{\beta}_{ts})$ it is not guaranteed that in practice the result will be positive definite. We therefore use a different method that guarantees positive definiteness (see Newey 1984). It can be shown that $\sqrt{T}(\hat{\beta}_{ml}-\hat{\beta}_{ts})=\sqrt{T}(\hat{\beta}_{ml}-\beta)-\sqrt{T}(\hat{\beta}_{ts}-\beta) \xrightarrow{d} N(0,F)$

where $F=[H_{ml}^{-1}G_{ml}H_{ml}^{-1'}+H_{ts}^{-1}G_{ts}H_{ts}^{-1'}-H_{ml}^{-1}G_{ml/ts}H_{ts}^{-1'}-H_{ts}^{-1}G_{ts/ml}H_{ml}^{-1'}]$

and

H_{ml} $=E[\partial^2 L/\partial\beta\partial\beta']$,the hessian matrix of the maximum likelihood estimator

G_{ml} $=E[(\partial L/\partial\beta)(\partial L/\partial\beta)']$,the cross product of the first derivatives of the ML estimator

G_{ts} =a block matrix $\begin{bmatrix} A & 0 \\ C & D \end{bmatrix}$ with

$A=E[\partial^2 L_1/\partial\beta_1\partial\beta_1']$, the hessian of the first step with respect to β_1

$C=E[\partial^2 L_2/\partial\beta_2\partial\beta_1']$, the derivative of the moment conditions in the second step with respect to β_1

$D=E[\partial^2 L_2/\partial\beta_2\partial\beta_2']$, the hessian of the second step with respect to β_2

G_{ts} $=E[f\ f']$ with f defined as the stacked vectors $(\partial L_1/\partial\beta_1)$ and $(\partial L_2/\partial\beta_2)$.

$G_{ml/ts}$ $=E[(\partial L/\partial\beta)f']$

$G_{ts/ml}$ $=E[f(\partial L/\partial\beta)']$

In the estimation of F, the expectations are replaced by their sample equivalents evaluated at the estimated $\hat{\beta}$.

Table 4C.1. Ordered probit estimates (ML std errors corrected, two step std errors not corrected)

MALES:	ML		two step		FEMALES: ML		two step	
	coeff	std. error	coeff	std. error	coeff	std. error	coeff	std. error
α_1	-0.159	0.218	-0.079	0.229	-1.961	0.281	-2.008	0.278
δ_2:cnst	-1.919	0.322	-1.843	0.318	-1.177	0.262	-1.137	0.261
young	0.007	0.013	0.000	0.013	0.007	0.011	0.007	0.011
prime	-0.050	0.015	-0.051	0.015	-0.039	0.012	-0.044	0.012
old	0.059	0.065	0.093	0.059	0.004	0.050	0.000	0.050
married	0.996	0.186	1.000	0.193	0.023	0.138	-0.014	0.137
net dsav pc /1000	-2.389	0.726	-2.179	0.718	-0.279	0.410	-0.311	0.416
ehat-nets	-0.100	0.727	-0.431	0.719	-0.137	0.410	-0.084	0.416
mar*age	-0.018	0.004	-0.017	0.005	0.000	0.003	0.000	0.003
age	0.093	0.014	0.095	0.013	0.033	0.011	0.033	0.011
age square /100	-0.085	0.016	-0.089	0.016	-0.020	0.012	-0.020	0.012
inter	0.084	0.063	0.075	0.062	-0.092	0.057	-0.089	0.058
medio	-0.047	0.062	-0.062	0.061	-0.393	0.057	-0.388	0.058
midtech	-0.262	0.154	-0.293	0.151	-0.981	0.133	-0.979	0.133
hightech	-0.199	0.182	-0.241	0.181	-1.077	0.231	-1.071	0.229
normal	-0.484	0.226	-0.531	0.223	-1.625	0.167	-1.626	0.167
university	-0.593	0.103	-0.644	0.100	-1.254	0.134	-1.253	0.134
other	-0.228	0.127	-0.219	0.132	0.016	0.065	0.013	0.065
yrs incompl educ	-0.033	0.019	-0.046	0.019	-0.079	0.014	-0.079	0.014
econ activity /10	-0.001	0.036	-0.010	0.036	0.070	0.028	0.066	0.028
unemployment *10	-0.308	0.143	-0.384	0.138	0.319	0.115	0.303	0.115
ethnic	0.096	0.048	0.087	0.047	0.321	0.039	0.318	0.039
δ_1:young	0.013	0.009	-0.003	0.011	0.043	0.013	0.045	0.014
prime	0.031	0.009	0.066	0.011	0.011	0.012	0.008	0.012
old	0.071	0.038	0.077	0.052	0.023	0.053	0.019	0.054
married	0.061	0.108	0.114	0.136	1.098	0.138	1.085	0.137
net dsav pc /1000	1.587	0.454	2.478	0.552	0.942	0.446	0.885	0.442
ehat-nets	-0.673	0.453	-2.183	0.551	-0.544	0.451	-0.555	0.446
mar*age	-0.008	0.003	-0.007	0.004	-0.011	0.004	-0.010	0.004
age	-0.035	0.010	-0.033	0.010	-0.124	0.013	-0.126	0.012
age square /100	0.072	0.012	0.065	0.013	0.181	0.016	0.184	0.016
inter	-0.060	0.052	-0.076	0.053	-0.139	0.074	-0.137	0.074
medio	-0.111	0.047	-0.154	0.048	-0.325	0.063	-0.319	0.063
midtech	-0.321	0.092	-0.380	0.096	-0.964	0.088	-0.958	0.088
hightech	-0.349	0.117	-0.447	0.117	-1.196	0.129	-1.188	0.129
normal	-0.998	0.106	-1.120	0.103	-1.977	0.078	-1.973	0.078
university	-0.598	0.063	-0.706	0.066	-1.424	0.082	-1.418	0.081

Table 4C.1. Ordered probit estimates continued (ML std errors corrected, two step std errors not corrected)

	coeff	std. error	coeff	std. error	coeff	std. error	coeff	std. error
other	-0.305	0.091	-0.236	0.095	0.118	0.091	0.118	0.091
yrs incompl educ	-0.037	0.013	-0.062	0.014	-0.109	0.017	-0.109	0.017
econ activity /10	0.017	0.027	-0.004	0.027	-0.017	0.032	-0.021	0.032
unemployment *10	0.012	0.105	-0.035	0.107	0.660	0.126	0.654	0.125
ethnic	0.158	0.036	0.154	0.037	0.256	0.049	0.253	0.049
β_1:cnst	0.042	0.174	0.317	0.252	-1.777	0.288	-1.892	0.301
age	0.026	0.008	0.019	0.009	0.081	0.012	0.083	0.013
age square /100	-0.007	0.010	0.005	0.013	-0.088	0.017	-0.092	0.018
inter	0.082	0.049	0.069	0.044	0.220	0.088	0.227	0.090
medio	0.245	0.041	0.225	0.038	0.523	0.075	0.539	0.076
midtech	0.440	0.078	0.385	0.073	0.813	0.104	0.854	0.100
hightech	0.582	0.100	0.532	0.084	1.028	0.141	1.079	0.124
normal	0.078	0.083	-0.050	0.095	0.972	0.121	1.045	0.116
university	0.827	0.048	0.748	0.061	1.311	0.110	1.370	0.106
other	0.211	0.067	0.146	0.073	-0.209	0.098	-0.214	0.118
yrs incompl educ	0.061	0.012	0.057	0.011	0.088	0.017	0.092	0.017
econ activity /10	0.172	0.023	0.180	0.020	0.176	0.032	0.176	0.030
unemployment *10	-0.664	0.090	-0.665	0.082	-0.704	0.121	-0.723	0.118
ethnic	-0.061	0.034	-0.037	0.031	-0.183	0.055	-0.194	0.050
β_2:cnst	0.221	0.278	0.327	0.252	0.273	0.331	0.169	0.331
age	-0.005	0.012	-0.017	0.013	0.009	0.014	0.013	0.014
age square /100	0.030	0.015	0.047	0.016	0.000	0.017	-0.005	0.017
inter	0.052	0.061	0.052	0.053	0.256	0.071	0.257	0.064
medio	0.284	0.054	0.291	0.048	0.360	0.064	0.364	0.060
midtech	0.377	0.136	0.384	0.120	0.566	0.121	0.589	0.136
hightech	0.371	0.169	0.363	0.144	0.813	0.245	0.844	0.231
normal	-0.419	0.166	-0.545	0.177	0.125	0.175	0.199	0.189
university	0.600	0.080	0.565	0.082	0.718	0.117	0.760	0.143
other	-0.404	0.106	-0.401	0.107	-0.042	0.074	-0.046	0.072
yrs incompl educ	0.046	0.018	0.037	0.016	-0.002	0.019	-0.001	0.018
econ activity /10	0.136	0.034	0.128	0.031	0.151	0.035	0.153	0.035
unemployment *10	-0.889	0.133	-0.890	0.119	-0.854	0.149	-0.866	0.141
ethnic	-0.073	0.045	-0.063	0.041	-0.145	0.048	-0.145	0.045
sigmaf	0.905	0.012			0.684	0.011		
sigmai	0.986	0.026			0.954	0.017		
sigf	0.704	0.023	0.949	0.149	-0.080	0.076	-0.140	0.075
sigi	0.653	0.040	0.765	0.090	0.268	0.062	0.216	0.069
Log Likelihood:11601 nr obs:6349					Log Likelihood:10160		nr obs:7293	

Table 4C.2. Multinomial logit estimates (ML std errors corrected, two step std errors not corrected)

MALES:	ML		two step		FEMALES ML		two step	
	coeff	std. error	coeff	std. error	coeff	std. error	coeff	std. error
δ_1:cnst	-1.079	0.519	-1.075	0.527	-4.269	0.547	-4.424	0.559
young	0.003	0.025	0.030	0.027	-0.101	0.026	-0.102	0.027
prime	-0.187	0.023	-0.221	0.025	-0.064	0.023	-0.059	0.025
old	0.006	0.116	-0.051	0.122	-0.030	0.102	-0.025	0.104
married	0.511	0.278	0.487	0.307	-2.436	0.271	-2.381	0.285
net dsav pc /1000	-6.247	1.124	-7.310	1.136	-2.525	0.875	-2.242	0.906
ehat-nets	2.871	1.119	4.687	1.144	1.349	0.888	1.030	0.920
mar*age	0.009	0.007	0.007	0.008	0.023	0.008	0.021	0.008
age	0.208	0.023	0.213	0.023	0.311	0.025	0.321	0.026
age square /100	-0.295	0.029	-0.297	0.030	-0.436	0.033	-0.450	0.034
inter	0.260	0.140	0.298	0.141	0.272	0.149	0.265	0.145
medio	0.138	0.122	0.223	0.123	0.500	0.126	0.484	0.122
midtech	0.359	0.229	0.397	0.232	1.527	0.164	1.527	0.161
hightech	0.394	0.266	0.466	0.271	1.926	0.236	1.927	0.225
normal	1.446	0.301	1.572	0.289	3.294	0.158	3.293	0.158
university	0.566	0.154	0.725	0.156	2.385	0.157	2.353	0.157
other	0.107	0.216	0.025	0.209	-0.221	0.193	-0.234	0.189
yrs incompl educ	0.032	0.033	0.061	0.034	0.208	0.031	0.206	0.033
econ activity /10	-0.019	0.067	-0.007	0.067	0.110	0.061	0.109	0.062
unemployment *10	-0.735	0.263	-0.759	0.275	-1.248	0.241	-1.287	0.240
ethnic	-0.167	0.092	-0.177	0.091	-0.339	0.097	-0.347	0.094
δ_2:cnst	-3.093	0.566	-3.043	0.579	-4.496	0.432	-4.446	0.428
young	0.024	0.026	0.027	0.029	-0.039	0.018	-0.049	0.020
prime	-0.172	0.026	-0.178	0.028	-0.081	0.019	-0.097	0.021
old	0.161	0.118	0.146	0.133	-0.013	0.090	-0.018	0.093
married	1.114	0.295	1.488	0.330	-0.937	0.213	-1.082	0.224
net dsav pc /1000	-4.621	1.186	-5.519	1.243	-2.269	0.753	-2.312	0.755
ehat-nets	3.207	1.185	2.679	1.245	-0.236	0.757	0.574	0.767
mar*age	-0.014	0.007	-0.019	0.008	0.005	0.005	0.006	0.005
age	0.265	0.025	0.256	0.025	0.239	0.018	0.248	0.018
age square /100	-0.306	0.030	-0.292	0.031	-0.281	0.022	-0.294	0.022
inter	0.205	0.143	0.244	0.145	-0.034	0.091	-0.020	0.090
medio	-0.052	0.127	-0.018	0.128	-0.352	0.084	-0.355	0.083
midtech	-0.282	0.263	-0.313	0.263	-0.532	0.173	-0.531	0.173
hightech	-0.326	0.308	-0.323	0.310	-0.458	0.299	-0.454	0.290
normal	-0.358	0.350	-0.451	0.345	-0.448	0.209	-0.461	0.209
university	-0.642	0.178	-0.662	0.177	-0.530	0.172	-0.531	0.175

Table 4C.2. Multinomial logit estimates continued (ML std errors corrected, two step std errors not corrected)

	coeff	std. error	coeff	std. error	coeff	std. error	coeff	std. error
other	-0.580	0.226	-0.581	0.227	-0.109	0.104	-0.135	0.104
yrs incompl educ	-0.035	0.036	-0.044	0.037	-0.052	0.026	-0.050	0.026
econ activity /10	-0.012	0.073	-0.024	0.073	0.154	0.049	0.157	0.048
unemployment *10	-1.153	0.285	-1.179	0.297	-0.226	0.195	-0.292	0.193
ethnic	0.139	0.098	0.129	0.097	0.431	0.065	0.437	0.065
β_1:cnst	-0.058	0.168	-0.757	0.195	-1.887	0.278	-1.911	0.298
age	0.031	0.008	0.052	0.008	0.083	0.012	0.084	0.013
age square /100	-0.013	0.010	-0.046	0.011	-0.092	0.016	-0.093	0.018
inter	0.082	0.048	0.093	0.044	0.227	0.088	0.228	0.090
medio	0.248	0.040	0.268	0.038	0.540	0.073	0.543	0.076
midtech	0.441	0.076	0.513	0.071	0.853	0.100	0.860	0.098
hightech	0.582	0.099	0.679	0.081	1.079	0.137	1.087	0.122
normal	0.106	0.080	0.372	0.071	1.044	0.112	1.054	0.112
university	0.854	0.046	1.002	0.047	1.371	0.102	1.381	0.104
other	0.211	0.064	0.290	0.070	-0.210	0.098	-0.211	0.118
yrs incompl educ	0.062	0.011	0.082	0.010	0.094	0.017	0.095	0.017
econ activity /10	0.169	0.022	0.176	0.020	0.177	0.032	0.178	0.030
unemployment *10	-0.676	0.087	-0.632	0.082	-0.727	0.121	-0.727	0.118
ethnic	-0.066	0.033	-0.111	0.029	-0.194	0.054	-0.197	0.050
β_2:cnst	-1.961	0.302	1.357	0.469	1.874	0.387	2.419	0.479
age	0.095	0.013	0.011	0.015	-0.024	0.015	-0.039	0.016
age square /100	-0.094	0.015	-0.008	0.016	0.034	0.018	0.050	0.019
inter	0.077	0.065	0.083	0.054	0.275	0.076	0.286	0.063
medio	0.204	0.058	0.369	0.052	0.529	0.069	0.563	0.065
midtech	0.215	0.137	0.679	0.132	1.001	0.131	1.089	0.144
hightech	0.143	0.167	0.684	0.159	1.375	0.251	1.471	0.236
normal	-0.610	0.166	0.649	0.227	1.208	0.161	1.363	0.200
university	0.227	0.091	1.159	0.140	1.407	0.127	1.474	0.150
other	-0.537	0.118	-0.065	0.120	-0.042	0.080	-0.041	0.071
yrs incompl educ	0.017	0.018	0.081	0.018	0.039	0.021	0.045	0.019
econ activity /10	0.118	0.036	0.126	0.031	0.117	0.038	0.113	0.036
unemployment *10	-1.218	0.145	-0.814	0.131	-0.904	0.157	-0.922	0.138
ethnic	-0.028	0.048	-0.228	0.047	-0.311	0.053	-0.350	0.053
sigmaf	0.878	0.011			0.689	0.013		
sigmai	1.110	0.033			1.138	0.039		
sig13	0.655	0.022	0.115	0.081	-0.138	0.068	-0.146	0.070
sig23	-0.923	0.048	0.741	0.204	0.803	0.070	0.988	0.138
Log Likelihood:11635 nr obs:6349					Log Likelihood:10120 nr obs:7293			

5. Sector Participation in Labour Supply Models: Preferences or Rationing?

5.1 Introduction

Empirical studies using the dual labour market approach for developing countries have focused on testing for labour market segmentation. A test for labour market segmentation usually works as follows: wages (or hourly earnings) of workers with a similar background are compared for the formal and informal sector. It is assumed that in selecting the preferred sector the potential workers consider the wage offers only. In the absence of rationing every individual would enter the sector offering the highest wage and, in equilibrium, wage offers would equalize. In the presence of rationing, however, some workers desiring a formal sector job (because the wage offer is higher) may not be able to obtain one and end up working in the informal sector. Thus, labour market segmentation implies that there are identical workers with different earnings in the two sectors. See Dickens and Lang (1992) for a survey of segmentation theory and evidence.

A shortcoming of the applied tests for labour market segmentation is that they have not been able to make a distinction between wage differentials resulting from individual preferences concerning non-wage job characteristics and those resulting from restrictions in mobility between sectors. While the latter can be regarded as an indicator of "true" labour market segmentation, compensating wage differentials resulting from preferences would remain existent, even in the absence of rationing. Compensating wage differentials follow from non-monetary returns to the job such as health insurance, utility associated with the workplace or job security. Also pure preferences, for example following from the social status attached to participation in a particular sector, could cause a wage differential between sectors. Gindling (1991) circumvented this problem by comparing wages within narrowly defined occupational groups, assuming that within these groups the non-monetary returns are equal across sectors.

In this chapter we will propose a model that allows us to test for labour market segmentation between the formal and informal sector in a developing country on the basis of cross-sectional data. The proposed test overcomes the shortcomings that were discussed in the previous paragraph. The model exploits the specific features of urban labour markets in developing countries, in particular the existence of a competitive informal sector. The model incorporates information on search, including on the job search, to identify rationed individuals. The organization of the chapter is as follows: in section 5.2 we discuss the underlying assumptions of the model. Especially the treatment of information on search and discouragement is a novelty of the model. On the basis of these assumptions we derive in section 5.3 a formal model of labour supply that deals simultaneously with wages, sector preferences and rationing. Section 5.4 describes the data on search activity and looks for empirical evidence of rationing by means of descriptive statistics. In section 5.5 the estimation strategy for the formal model is

discussed. In section 5.6 the estimated coefficients are presented, a test for labour market segmentation is derived and some simulations are presented that could be interesting from a policy point of view. Section 5.7, finally, concludes.

5.2 Rationing, Search and Discouraged Workers

A labour supply model in which individuals face no entry barriers for any of the sectors typically comprises two parts: a sector selection part and a wage determination part. In the sector selection part, the sector in which the individual decides to participate is explained. In a Heckman (1974) type model the choice would be between working and not working, but in principle more than two sectors can be allowed for. Sector choice is explained by the wage offers in each sector (for non-participation the value of home production) and individual preferences associated with participation in a sector. In the wage determination part, wage offers in each sector are related to individual characteristics and local labour market conditions. The wage equations have to be estimated jointly with the sector selection section because of non-random selection into sectors. For the Bolivian setting, models of this kind were analyzed in the previous chapter. In this chapter the sector selection section will be replaced by a sector preference section and a rationing section.

We want to develop a labour supply model that explicitly considers rationing and that can be estimated on cross-sectional household survey data stemming from urban areas in a developing country. To this end, some assumptions considering the sectors and the treatment of rationing have to be made. These are:
-The informal sector is a competitive, free entry sector.
-Rationing is indicated by search or discouragement.

The first assumption is rather standard. It is often used as a way to define the informal sector. It is also mentioned in a recent review of the informal sector studies by the OECD (Lubell 1990). Firms operating in the informal sector are small and require low capital investments. Anyone wishing to work in the informal sector can do so immediately. As a consequence, all rationing takes place for formal sector jobs only.

The second assumption requires more discussion. If an individual searches, this indicates that he (or she) is not content with his (or her) present situation. If the individual were not rationed he (or she) would change to the preferred situation. Search indicates that the person is not able to do so immediately and is rationed. There may, however, be individuals who are rationed but do not search. They may have searched in the past but were unsuccessful in their quest. In our analysis we treat these discouraged persons (those who report not to be searching for work because no work is available) identical to the searchers. Finally, it is also possible that an individual searches but is not rationed. In that case the individual just wants to determine his or her market value by looking what is available. The model does not allow for this: the individual is assumed to have perfect information on the distribution of potential earnings in all sectors.

One way to interpret discouragement is the presence of search costs: some people will not search because the costs of search exceed the expected gains. These people are therefore rationed in the sense that their long run preferred status is different from the actual state, but they do not search. By treating them as rationed, search costs do not play a role. The model thus does not explain the search-no search decision, but only the preferred state, ignoring search costs. In principle, the distinction between discouraged and searching individuals could be useful to identify search costs but, given the nature of the cross section data, without information on how long it takes to find a job, this approach does not seem to be feasible. In the sequel, the term searchers will be used for both searching and discouraged individuals.

Under the assumptions stated above, rationing takes place for formal sector jobs only and individuals rationed for formal sector jobs are identified by their search activity or because they are discouraged. The sector in which an individual is observed is influenced by wage offers, preferences and rationing. As yet, the model does not allow for individuals to search while they are working in the formal sector. Since they are working in the formal sector, they cannot be rationed for this sector at the same time. Such observations, however, do appear in the dataset. To allow for this situation we introduce two types of formal sector jobs: "good" and "bad" jobs. Good jobs have the advantage over "bad" jobs in that they offer a higher wage. In other respects, such as the non-monetary returns, the two type of jobs are assumed to be the same[24]. Thus, any individual will always prefer a "good" formal sector job to a "bad" one. Those who work in the formal sector and search are assumed to have a "bad" formal sector job and to be looking for a "good" one. Those who do not search have a "good" formal sector job. Search or discouragement thus identifies the difference between "good" and "bad" in the data. In this setting we assume that rationing takes place for "good" formal sector jobs only. The nature of the (cross section) data does not allow for a continuum of bad-good job offers.

The model described above allows us to specify for each individual the preferred sector, the actual sector and whether the individual is rationed or not. All the possible cases are given in table 5.1. Entry into the "good" formal sector is rationed, all other sectors are assumed to operate competitively. Non-workers, by definition, include both voluntary and involuntary unemployed. For formal and informal sector workers we define search as looking for another job replacing the current one. One could think of these workers as underemployed. As hours of work are not considered, those individuals who report to be looking for additional work are categorized as non-searchers. For non-participants and workers in the informal sector that do not search we know that their actual status is the preferred one. For them we do not know whether they are rationed for "good" formal

[24]In principle, the model can be adjusted to allow for the utility of "good" and "bad" formal jobs to differ with taste shifters as well. However, since both "good" and "bad" jobs are similar in the sense that they are both salaried work we do not allow for this in the estimation. Only the difference in non-monetary returns between the formal and informal sector is analyzed.

sector jobs. They can directly enter into the preferred sector and rationing for the formal sector is irrelevant to their actual status.

Table 5.1 Relationship between observed, preferred sector and rationing

Observed Sector	searching	Preferred Sector	Actual Sector	Rationed
formal sector	yes	(good) formal	bad formal	yes
	no	(good) formal	good formal	no
informal sector	yes	(good) formal	informal	yes
	no	informal	informal	?
non-participation	yes	(good) formal	non-participation	yes
	no	non-participation	non-participation	?

Identification of the effects of preferences versus rationing is achieved through incorporation of search information in the model. The coefficients in the rationing section are identified on the basis of the search information of formal sector workers. The sector participation decision of non-searchers in combination with their observed wages identify the coefficients in the preference section of the model. Only the difference in preferences attached to participation in the informal versus non-participation and participation in the formal versus non-participation can be identified. The difference in the wage offer function between good and bad formal sector jobs is identified on the basis of the observed wages of searchers and non-searchers in the formal sector and on the assumption that informal sector searchers prefer a good formal sector job to their current one while this is not the case for a bad formal sector job.

5.3 The Model

We model the case where monetary and non-monetary returns are perfect substitutes. Non-monetary returns are defined as the monetary equivalent of all non-monetary benefits associated with participation in a sector. The individual is assumed to maximize a utility function depending on leisure (T-h) and consumption (C^*), including non-monetary consumption. Within a two-stage budgeting framework, consumption in each period has to be less than or equal to total earnings plus net dissavings (See section 4.2 for a discussion on net dissavings). The budget constraint is assumed linear because income taxes play a minor role in Bolivia (cf. chapter 2 page 12). The maximization problem can be written as follows

$$\max\ u(T-h,C^*) \qquad s.t.\quad C^* \leq w^* h + Y \tag{5.1}$$

with

w^* = the hourly earnings, including non-monetary,
Y = net dissavings,

h = hours worked,
T = time endowment.

Hourly earnings are a function of monetary and non-monetary earnings. Non-monetary earnings vary with the sector in which the individual participates.

$$w^* = w_j * \lambda_j \qquad \begin{matrix} j=1 \textit{ if in formal sector} \\ j=2 \textit{ if in informal sector} \end{matrix} \tag{5.2}$$

with
w_j = the hourly wage in sector j,
λ_j-1 = non-monetary returns in each sector modelled as a fraction of the wage.
Under the assumption that hourly wage and non-monetary returns do not depend on hours worked and that the individual is not restricted in the number of working hours it follows that the individual will participate if the hourly earnings w^* in at least one sector exceed the reservation wage, w_R. The reservation wage equals the marginal rate of substitution between leisure and consumption evaluated at zero hours of work. If participating, the individual will participate in the sector offering the highest w^*. In this chapter we focus on the participation decision and ignore hours worked. The complete labour supply decision is considered in the next chapter.

Wage offers. Wage offers in the formal and informal sector are modelled as a function of individual specific variables and regional specific variables describing the local labour market conditions. In the formal sector two wage offers exist, a "good" and a "bad" one. They differ in that the wage offer for the "good" one is higher than for the "bad" one. In other respects, such as the non-monetary returns to the job, the two types of jobs are identical. Bad formal sector jobs can be regarded as underpaid. Any individual will thus prefer a "good" formal sector job to a "bad" formal sector job. We start by formulating the "good" formal sector wage. In the informal sector (j=2) no distinction between good and bad exists. The potential log hourly wage of individual i in sector j is modelled as

$$\ln(w_{ij}^g) = X_i\beta_j + \eta_{ij} \qquad j=1,2 \tag{5.3}$$

with
w_{ij}^g = wage offer of a "good" formal sector job,
X_i = a vector of explanatory variables influencing returns,
η_{ij} = an error term, representing unobserved factors influencing productivity.

Bad formal sector jobs offer a lower wage than good formal sector jobs. This is imposed in the estimation by including an exponential term on the difference in the wage offer between "good" and "bad" formal sector jobs. The difference in the wage offer is made dependent on individual characteristics and local labour market conditions. The potential log hourly wage for a "bad" formal sector job equals

$$\ln(w_{i1}^{b}) = \ln(w_{i1}^{g}) - \exp(X_i\tau) \tag{5.4}$$

with
w_{i1}^{b} = wage offer of a "bad" formal sector job.

Preferences. Because both non-monetary returns and the reservation wage are unobserved we only model the difference between the two. For individual i the model is

$$\ln(\lambda_{ij}) - \ln(w_{iR}) = Z_i\delta_j + \epsilon_{ij} \qquad j=1,2 \tag{5.5}$$

with
Z_i = a vector of explanatory variables influencing preferences for sectors,
ϵ_{ij} = an error term, representing any unobserved taste shifters influencing non-monetary returns and/or the reservation wage.

Z_i contains individual and household specific variables. Z_i includes a measure of net dissavings because it influences the reservation wage. Characteristics of the observed jobs are not included as they are not constant across sectors.

Rationing. Barriers to entry in the "good" formal sector are modelled in the rationing equation. Rationing for "good" formal sector jobs is assumed to be a discrete zero/one variable and to depend on individual and labour market characteristics influencing the demand for formal sector labour. We model the probability that an individual is rationed by

$$R_i^* = R_i\alpha + \epsilon_{ri}^* \qquad \textit{with} \qquad \begin{matrix} \textit{rationed if } R_i^* > 0\,, \\ \textit{not rationed if } R_i^* < 0 \end{matrix} \tag{5.6}$$

with
R_i^* = a latent variable indicating whether individual i is rationed,
R_i = Explanatory variables influencing rationing,
ϵ_{ri}^* = an error term, representing "luck" in entering into the good formal sector.

Error structure $\epsilon_1, \epsilon_2, \eta_1, \eta_2$ and ϵ_r^* are assumed to be i.i.d. with a multivariate normal distribution with mean zero and covariance matrix Σ.

$$\epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \eta_1 \\ \eta_2 \\ \epsilon_r^* \end{pmatrix} \sim N(0, \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 \\ & \sigma_2^2 & 0 & 0 & 0 \\ & & \sigma_f^2 & & \sigma_{fr} \\ & & & \sigma_i^2 & \sigma_{ir} \\ & & & & 1 \end{pmatrix}) \tag{5.7}$$

A correlation between the error terms in the wage equations and the rationing equation is allowed for. The correlation between the errors of the wage equation in the formal and informal sector is not identified because we observe an individual working in at most one sector. The variances in the preference section equations (σ_1^2 and σ_2^2) are identified because of exclusion restrictions in Z and X. Rationing is modelled in a probit type specification and therefore the variance of ϵ_r^* is normalized to one. The errors in the rationing equation and the preference section equations are assumed to be independent. However, some correlation is introduced indirectly as omitted variables appearing in the wage offer equations are allowed to influence both the probability of rationing as well as the probability of participation in sectors.

5.4 Data

In chapters 5 and 6 add a capital asset condition to the sector assignment rule. By doing this, we want to capture credit constraints that may exist for some jobs in self-employment. From now on we classify self-employed workers as formal sector workers if their household business assets are greater than 15.000 bolivianos (+/− u$ 5500). Business assets include property of land, car for business use and telephone. The other self-employed (about 95 percent) remain categorized as informal sector workers. As a result 1.6 percent of the males (100 obs) and 0.8 percent of the females (55 obs) change sector definition from informal to formal. The idea behind the definition change is that businesses requiring high assets are not in reach for every individual. Potential entrepreneurs in the informal sector face credit constraints because their activities cannot be well monitored by the lending institution. It implies that jobs in these businesses are rationed. Defining them as informal is incompatible with the assumption that informal sector jobs are not rationed, which is an explicit assumption in this chapter. A revised table of descriptive statistics is in appendix 5A. The change in most of the variables is very small. The most noticeable change is that average formal sector wages now exceed informal sector wages for both males and females. The difference, however, remains insignificant.

As indicated in section 5.2, we treat active searchers and discouraged individuals as rationed in the analysis. In the 1989 survey, both non-participants as well as workers were asked whether or not they engaged in active search for work. In case they did not, they were asked why they did not search. We have classified as discouraged those who answered to the latter "No hay trabajo" (there is no work). This answer expresses a belief that it would be very hard to get a formal sector job or, in other words, that search costs are too high. Unfortunately the answer "No vale la peina" (it is not worth the trouble) was not included in the answer list. This answer would be closer related to our view of discouragement. The questions relating to search in the 1989 survey are

copied in box 5.1[25]. Between brackets are the relative frequencies for all individuals between 18 and 65 in the survey.

For non-workers:	For workers:
¿Buscó Ud. la semana pasada trabajo?	¿Durante la semana pasada buscó Ud. alguna otra ocupación?
1. si (12)	1. si (8)
2. no (88)	2. no (92)
¿Por qué no buscó?	¿Por qué no buscó otro trabajo?
1. licencia, vacacion (3)	1. no alcanza el tiempo (66)
2. realizar labores de casa (54)	2. gana suficiente (5)
3. ser estudiante (24)	3. no hay trabajo (17)
4. ser rentista (4)	4. otra (12)
5. no hay trabajo (3)	
6. salud (6)	
7. otra razon (6)	

Box 5.1 Questions relating to search in EIH 89 (relative frequencies for individuals between 18 and 65 in survey between brackets).

Descriptive statistics on search behaviour are given in table 5.2. The pattern is very different for males and females. For males, most of the search takes place in the non-workers group. 53 percent of males who do not work are involuntary unemployed. For females, the amount of search is much lower for non-participants. Staying at home is in most of the cases the preferred status. For both males and females, on the job search activity is higher in the informal sector than in the formal. This suggests that the informal sector may function as a buffer, for those who seek a good job in the formal sector. Getting out of this buffer sector is not that easy, however. The percentage of discouraged searchers is the highest in the informal sector. This high rate of discouragement could be caused by high costs of on the job search in the informal sector or by an adverse selection and/or stigma effect. In the estimation active searchers and discouraged workers are treated identically. In our model, formal sector workers are assumed to be searching because they have an underpaid, "bad" job. To get a rough idea whether this is indeed the case we did a log hourly wage regression for all formal sector workers including a dummy for search activity in the set of explanatory variables of the wage equation. For both males and females, the estimated coefficient came out negative, as expected. For males the estimated coefficient was -0.23 with a t statistic of -7.4 and

[25]Translation: For non workers: Did you look for work in the past week? (1) yes, (2) no. Why did you not search? (1) vacation, (2) engaged in household work, (3) student, (4) person of private means, (5) there is no work, (6) health, (7) other reason. For workers: During the past week did you look for any other job? (1) yes, (2) no. Why did you not look for other work? (1) do not have time, (2) earn enough, (3) there is no work, (4) other.

for females -0.13 with a t-statistic of -2.5[26]. This result has also been found for developed countries (see Devine and Kiefer 1991, p. 238).

Table 5.2. Search activity by sector and sex. (percentages)

	males			females		
	formal	informal	not-working	formal	informal	not-working
active searching	6.3	6.6	52.9	2.0	3.5	8.8
discouraged	13.4	22.7	10.9	11.9	14.7	2.4

5.5 Estimation Method

Estimation of the model has been achieved by smooth simulated maximum likelihood (SSML). Evaluation of the likelihood function involves the computation of multiple integrals. Since these integrals are of a dimension higher than two, numerical integration requires an infeasible amount of computer time. Lerman and Manski (1981) proposed a frequency simulator to approximate the probabilities that appear in the likelihood function. Their simulator involves taking multiple draws from the distribution of the error terms. It has the disadvantage that a large number of draws is needed to obtain precise estimates. Moreover, the simulated likelihood function is not a smooth function of the unknown parameters. Standard optimization routines (for example, Newton-Raphson), which require differentiability of the approximated likelihood cannot be used. We follow a method that is similar to that proposed by Stern (1992). The resulting simulated likelihood function is a smooth function of the unknown parameters of the model. The estimates obtained by maximizing the simulated log likelihood function are consistent if the number of draws per observation tends to infinity with the number of observations. Moreover, provided that draws for different individuals are independent, ML and SSML will be asymptotically equivalent as $H/\sqrt{N}\to\infty$, where H is the number of draws per individual and N is the number of observations (cf. Gourieroux and Monfort, 1990)

Define $\theta=(\delta_1,\delta_2,\beta_1,\beta_2,\tau,\alpha)$. The likelihood function can be written as

$$L = \prod_{i=1}^{n} Pr[c_{1j}(y_i,Q_i,\theta)<\tilde{u}_i<c_{2j}(y_i,Q_i,\theta)\,|\,Q_i,\zeta_i]\; f_j(\zeta_i|Q_i) \tag{5.8}$$

with

[26]Because of the endogeneity of the search dummy this regression is not valid in the framework as a whole. The results are given for descriptive purposes only. However, the significance of the results does not change if one instruments for search (Heckman 1978) using taste shifters as instruments. The estimated coefficient on the predicted probability of search remains negative. For males the estimated coefficient is -1.16 with a t-statistic of 3.7 and for females -1.21 with a t-statistic of 2.6.

$\tilde{u}_i = A_j \epsilon_i$
$\zeta_i = B_j \epsilon_i$
y_i, Q_i = vectors containing dependent and independent variables

The contribution to the likelihood function of one individual can be written as a conditional probability times a density. The calculation of the conditional probability involves the evaluation of multiple integrals. Given θ, y_i and X_i, the value of ζ_i can be calculated. The transformed error terms, ζ and $(\tilde{u} \mid \zeta)$, follow a normal distribution $N(0,B_j\Sigma B_j')$ and $N(\mu,\Omega)$ with $\mu=A_j\Sigma B_j'(B_j\Sigma B_j')^{-1}\zeta$ and $\Omega=A_j\Sigma A_j'-A_j\Sigma B_j'(B_j\Sigma B_j')^{-1}B_j\Sigma A_j'$. The specification of the matrices A_j and B_j and the functions c_{1j} and c_{2j} depend on the regime (j) the individual is in. For example, in the case the individual is working in the informal sector and simultaneously looking for a job, A, B, c_1 and c_2 are defined by

$$A = \begin{bmatrix} -1 & 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \qquad B = [0\ 0\ 0\ 1\ 0] \tag{5.9}$$

$$c_1 = \begin{bmatrix} Z_i(\delta_1-\delta_2)+X_i\beta_1-\exp(X_i\tau)-w_{i2} \\ -\infty \\ -\infty \end{bmatrix} \quad c_2 = \begin{bmatrix} Z_i(\delta_1-\delta_2)+X_i\beta_1-w_{i2} \\ Z_i\delta_2+w_{i2} \\ R_i\alpha \end{bmatrix} \tag{5.10}$$

In this example, $\zeta_i=\eta_{i2}$ equals

$$\zeta_i=w_{i2}-X_i\beta_2 \tag{5.11}$$

The SSML estimator obtains an unbiased estimate of the probability in (5.8) by means of simulation. The idea is to decompose $\tilde{u}$ into two components, each normally distributed. The first is chosen such that it has a relatively simple covariance structure. The other is simulated. Let

$$\tilde{u}=u_1+u_2 \tag{5.12}$$

The Probability in (5.8) can now be written as

$$\begin{aligned} Pr[c_1<\tilde{u}<c_2 \mid X,\zeta] &= Pr[c_1-u_2<u_1<c_2-u_2 \mid X,\zeta] \\ &= E_{u_2}\{Pr[c_1-u_2<u_1<c_2-u_2 \mid X,\zeta,u_2] \mid X,\zeta\} \\ &= \int Pr[c_1-u_2<u_1<c_2-u_2 \mid X,\zeta,u_2]\ f(u_2 \mid X,\zeta)\ du_2 \end{aligned} \tag{5.13}$$

where the subscript i is suppressed. This is simulated by

$$\frac{1}{H}\sum_{k=1}^{H} Pr[c_1-u_{2k}<u_1<c_2-u_{2k}|X,\zeta,u_{2k}] \tag{5.14}$$

where the u_{2k} are drawn independently from $f(u_2 \mid X,\zeta)$, the conditional density of u_2, with H replications.

In the decomposition of $\tilde{u}$ we exploit the restrictions on the covariance matrix Σ. The decomposition varies with the regime the individual is in. We include in u_2, η_1 and/or η_2, if these are included in $\tilde{u}$. In the example stated above $u_2=[\eta_1,0,0]'$ and $\eta_1 \mid X,\zeta \sim N(0,\sigma_f^2)$. ζ does not enter here because the errors in the two wage equations are assumed independent. The distribution of $(u_1|X,\zeta,u_{2k})$ is normal $N(\mu_k,\Omega)$ with $\mu_k=A\Sigma B'(B\Sigma B')^{-1}(\zeta\ u_{2k})'$ and $\Omega=A\Sigma A'-A\Sigma B'(B\Sigma B')^{-1}B\Sigma A'$ where A and B are defined as

$$A=\begin{bmatrix}-1 & 1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & -1\end{bmatrix} \qquad B=\begin{bmatrix}0 & 0 & 0 & 1 & 0\\ 0 & 0 & 1 & 0 & 0\end{bmatrix} \tag{5.15}$$

In Stern (1992) the error terms u_1 and u_2 are assumed independent and the covariance matrix of u_1 is diagonal. In our method the error terms are dependent and we work with conditional distributions. The covariance structure of u_1 includes one non-zero covariance element. Computation of the probability remains feasible as this involves at most evaluating a bivariate cumulative probability. In the cases where it is unknown whether the individual is rationed, the contribution to the likelihood is calculated without simulation.

5.6 Estimation Results

The presented estimates coefficients are those obtained with H, the number of draws per observation equal to 30. The draws are different and independent for each observation. We experimented with taking different numbers of draws. A change from 10 to 30 draws made only a small difference in the estimation results. The reported estimates should thus be close to the exact ML estimates. The estimated coefficients are reported in tables 5.A.2. to 5.A.5. in the appendix 5A. Below we will discuss the results.

The estimated coefficients for the preference section of the model are presented in table 5.A.2. For both males and females the effect of net dissaving on the preference of working is negative. The hypothesis that net dissavings can be treated as exogenous is rejected for both sexes. For a discussion of why we used this specific other income

variable and how we corrected for potential endogeneity, see chapter 4[27]. The presence of other prime age individuals (PRIME) in the household significantly reduces the preference of working for both males and females. The effect is the strongest for males. The presence of young children (YOUNG) significantly reduces the preference of working for females. For males, formal sector preference increases. The model allows the effect of being married to be different for different age groups. In table 5.3 we report the partial derivatives of the probabilities that a sector is preferred with respect to MARRIED for different age groups and sexes. For males, MARRIED has a positive effect on the probability of working for every age group. This effect diminishes as the individual gets older. For females the effect is opposite, MARRIED increases the probability that non-participation is the preferred sector. Especially when females approach 30, being married drives them out of formal employment into non-participation.

Table 5.3. Partial effect of marriage on the probabilities that sectors are preferred (percentages, calculated at mean values)

	males			females		
age	20	30	40	20	30	40
formal	3.54	2.28	2.13	-5.54	-24.8	-30.8
informal	2.40	4.92	-0.09	-0.61	-5.76	-3.50
not working	-5.94	-7.20	-2.04	6.15	30.52	34.32

The estimated coefficients on the $Z\delta$ term in the preference equations can be used to get an idea of how the non-monetary returns compare in the formal and informal sector. For males, at mean values, non-monetary returns are higher in the formal sector than in the informal sector. To the make probabilities of preferring the formal and informal sector equal, a formal sector wage offer of only 0.46 (std err: 0.04) times that of the informal sector is needed. The additional security and benefits may cause males to value participation in the formal sector higher than in the informal sector. For females the opposite holds, there a formal wage offer of 1.18 (std err: 0.11) times that of the informal sector is needed to equalize probabilities.

To obtain insight in the effects of wages on the sector preferences, we have drawn figures 5.1 and 5.2. The figures show the probability a sector is preferred as a function of the wage offer. The figures are ceteris paribus, calculated at the mean values of the exogenous variables. The wage offer in one sector is varied, the wage offer in the other sector is held constant at the sample mean. At the mean value the log wage offers are for

[27]The standard errors of the ML estimates should be corrected for the error due to replacing the parameters of the instrumenting equation in the correction term by their estimates. The results in the previous chapter however showed that the effect of the correction is quite small. Therefore, for computational convenience, we have not corrected the standard errors here.

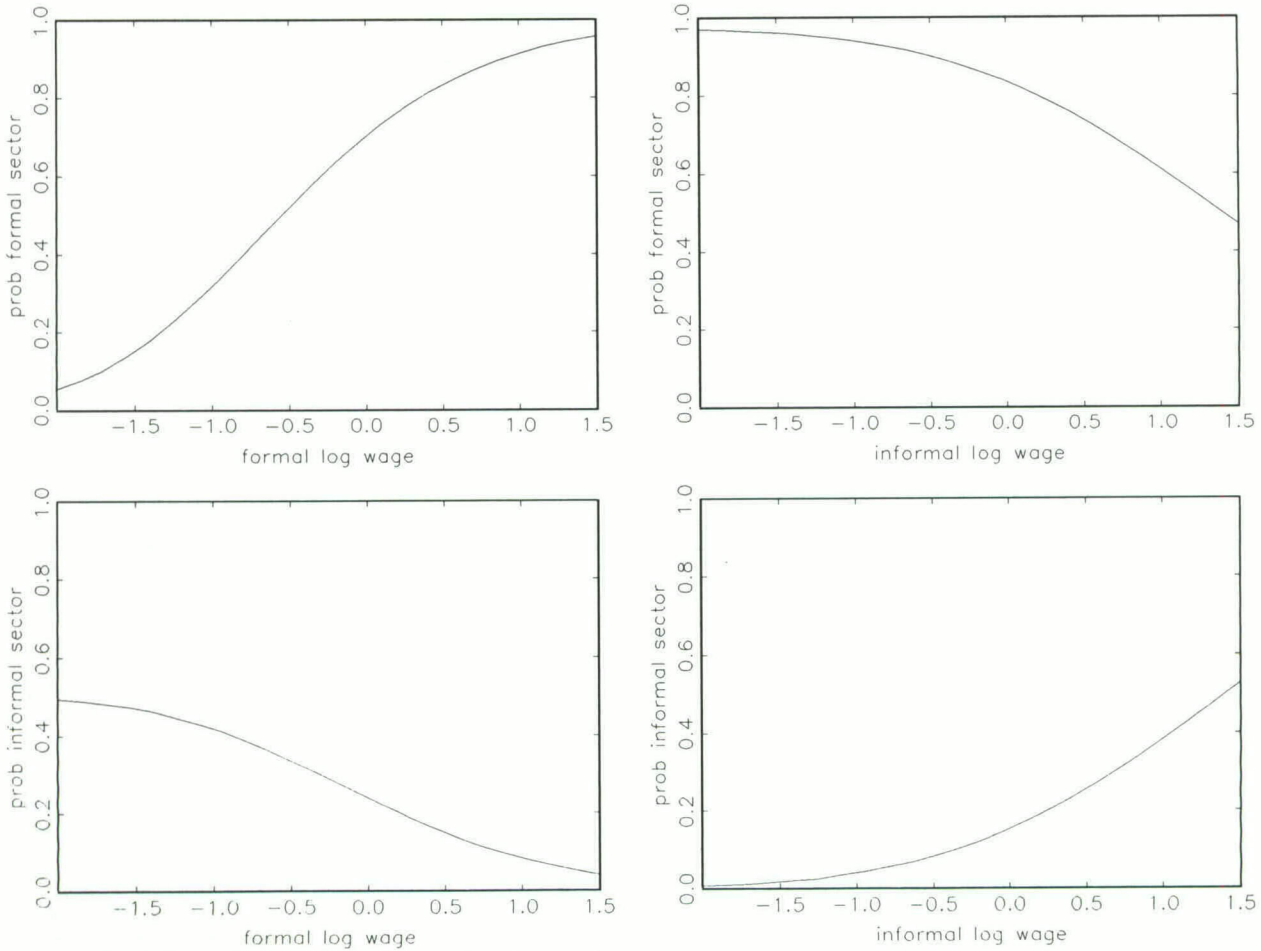

Figure 5.1 Effect of wage offers on preferred sector for males.

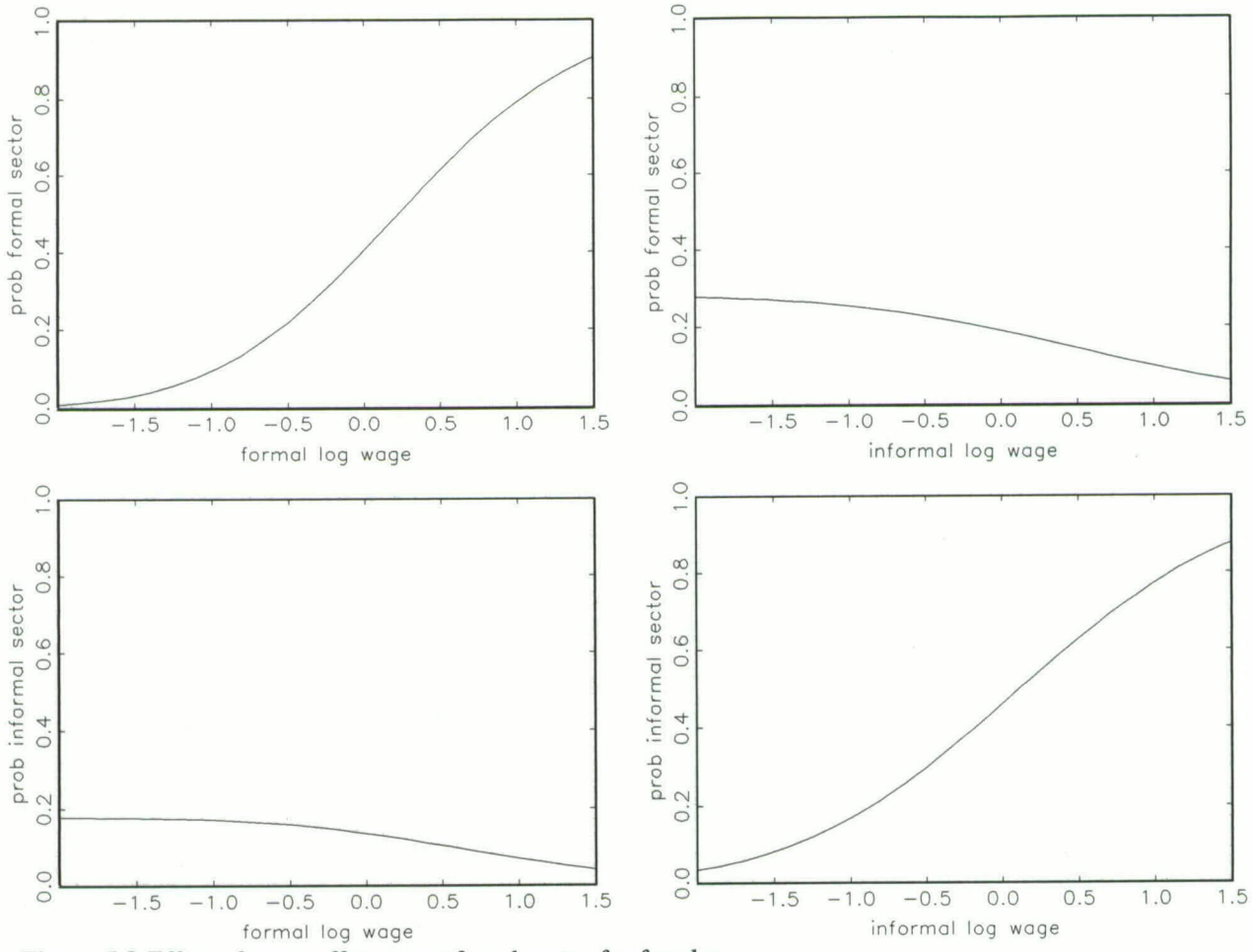

Figure 5.2 Effect of wage offers on preferred sector for females.

males (females, in parentheses) respectively: 0.62 (-0.38) in the formal sector and -0.14 (-1.06) in the informal sector. For males, the effect of a change in the wage offer is mostly a substitution effect between sectors. Since not-working is rarely the preferred status, an increase in the wage offer in one sector reduces the probability that the other sector is the preferred one. Only for very low formal sector wages, we find a probability that non-participation is preferred exceeding 0.1. For females the effects are different. For them an increase in the wage offer increases the probability of preference for that sector at the cost of not working. The own wage elasticities (relative change of probability of participation over a relative change in the wage offer) are 0.21 (1.35) for the formal sector and 1.19 (1.34) for the informal sector. The cross elasticities are -0.17 (-0.15) for the effect of the informal wage offer on the participation in the formal sector and -1.12 (-0.26) for the reverse.

The estimates for the wage equations are presented in table 5.A.3. The specification is the same as in chapter 4. Differences in the estimates may be due to the different definition of sectors and/or the change in selection bias correction. The estimates for the constant terms are quite different, most likely due to the change in selection bias correction (cf. chapter 4). Ceteris paribus, males reach their highest earnings at 45 years in the formal sector and at 48 years in the informal sector. For females the peaks are at 40 and 46 respectively. The effect of education is roughly similar in the two models. For males, returns to education are higher in the formal sector than in the informal sector. For females, education only has positive returns in the formal sector. In the informal sector only intermediate education has positive return (contrary to chapter 4). For some categories, such as "normal" even a significant negative return is estimated. Returns to education in the formal sector are higher for females than for males. For both males and females, the estimates are conform the segmented labour market hypothesis. We do allow the error terms of the wage equations to be correlated with the rationing equations. Whether such a correlation should be included for the informal sector is not obvious. In this sector it is less likely that there are unobserved factors which are used as rationing devices that influence productivity in the informal sector. For males, this correlation is not significantly different from zero at the 5 percent level. For females it is. The correlation between the formal sector wage equation and the rationing equation is positive and significant for both sexes. This indicates that formal sector workers (good and bad) who are overpaid (as a result of strong collective bargaining, for example) have a higher probability of being rationed.

The local unemployment rate has a significant negative effect on earnings in the informal sector. The effect is the strongest for males. In the formal sector, both for males and for females, no significant effect of unemployment can be detected (contrary to chapter 4). Belonging to an ethnic group significantly lowers the returns in the formal sector for males. For females, a positive effect of ethnicity is found in the informal sector (contrary to chapter 4). A possible explanation is that informal sector activity of females is largely concentrated in the commerce and restaurants sector (82 percent). In that sector a positive discrimination of ethnicity for females may exist.

Table 5.A.4 presents the estimates for τ, the coefficient in equation (5.4) that relates the good and bad formal wage offer. For males the difference is the greatest for higher technical education and university level training. For normal education the wage difference is relatively small. For females the effects are smaller. Only the local unemployment rate and the normal and university education variables have a significant effect. This suggests that underpayment is not correlated with observed individual characteristics for females. Their search activity in the (bad) formal sector may be more correlated with other, non-wage job characteristics. A likelihood ratio test rejected the null hypothesis that the wage differential, τ, does not depend on covariates for both sexes at the 5 percent level.[28]

Table 5.A.5. presents the estimates for the rationing equation for good formal sector jobs. A higher value of the left hand side variable is associated with an increased probability of rationing. The probability of rationing decreases with the level of education. This result holds for both males and females. This suggests that formal education is used as a screening device for formal sector jobs. A higher local unemployment rate increases the rationing level, as expected. Belonging to an ethnic group significantly increases rationing for females, for males the effect is much smaller. We experimented with incorporating information on migration into the rationing equation. For 2 out of 8 regions we have information about how long ago the individual moved to the place where he or she currently lives. In a Harris-Todaro type migration model, migrants work for a while in the informal sector while they queue up for formal sector jobs (cf section 3.2). Therefore we would expect the probability of rationing to increase if the individual moved in more recently. We could not find any significant effect of these variables. Because we only have the migration information for two capital cities we decided to present the results without the migration variables only.

Rationing can be viewed as an indicator for labour market segmentation. As is clear from the estimates, the degree of rationing varies strongly over individuals with different observed characteristics. It is therefore not possible to give a simple yes/no answer to whether there is labour market segmentation between the formal and informal sector in Bolivia. In table 5.4. we have calculated the probability of rationing for some standard individuals. For all individuals the probability of rationing is substantial, suggesting that the hypothesis of no labour market segmentation is rejected. The conclusions need not be that strong, however. If one believes in a natural rate of (frictional) search, one would test whether the predicted probability is significantly different from the natural search rate. The probabilities and their standard errors suggest that these frictional search rates

[28]As noted, it can be tested whether "good" and "bad" formal jobs differ with taste shifters (Z) as well. An LM test rejected the null hypothesis that good and bad formal jobs only differ in the wage offer as is assumed in the model at the 5% level. Estimating the model, allowing for the effect of taste shifters, causes additional complexities as the likelihood may not be defined for some individuals, since the utility of bad jobs might exceed that of good formal sector jobs.

are quite large. Note that labour market segmentation as defined in this model only refers to the rationing of "good" formal sector jobs.

Table 5.4. Predicted probability of rationing for some specific individuals (percentages, standard errors in parentheses, base: age=30, missed years=0, econ active and unempl at mean value)

		males	females
education	ethnic		
inter	no	33.2	34.1
		(1.98)	(3.88)
hightech	no	27.2	18.03
		(4.12)	(4.23)
university	no	20.0	21.3
		(1.66)	(2.62)
inter	yes	37.4	44.3
		(2.25)	(4.75)

The goodness of fit of the model can be examined by comparing the predicted probabilities of participation with the observed outcomes. The first 4 columns of tables 5.5 and 5.6 present the results for males and females. Predicted probabilities were obtained by simulation, taking multiple (50) draws from the estimated error distribution for each individual. The preferred status can be obtained by solely taking into consideration the predicted utilities of participation in a sector, setting the probability of rationing to zero. In general the predicted numbers are relatively close to the observed. For males, the model underestimates the number of individuals whose preferred sector is the formal sector, at the cost of the number of individuals for whom not working is the preferred status. For females, the number of individuals who prefer not to work is underestimated. Overall, the fit is reasonable.

Table 5.5. Males: Goodness of fit and simulation (number of individuals)

	Observed		Predicted		Unempl +5 percent	
	Actual	Preferred	Actual	Preferred	Actual	Preferred
Formal good	2973	4783	2782	4625	2088	4921
Formal bad	732		936		1275	
Informal	1763	1247	1780	1264	1462	837
Not working	881	319	850	460	1523	591

Table 5.6. Females: Goodness of fit and simulation (number of individuals)

	Observed		Predicted		Unempl +5 percent	
	Actual	Preferred	Actual	Preferred	Actual	Preferred
Formal good	1265	2280	1277	2281	765	2210
Formal bad	229		264		202	
Informal	1917	1568	1859	1674	1532	1324
Not working	3882	3445	3893	3338	4794	3759

The last two columns in table 5.5 and 5.6 present the results of a simulation exercise. We have simulated the effect of an exogenous increase in the rate of unemployment by 5 percent on sector participation and preferences. The results should be compared with the predictions from the base case (column 3 and 4). The effect of an increase in unemployment works in several ways. First, it has a negative effect on earnings in both sectors, thereby making non-participation more attractive. Second, the probability of rationing for "good" formal sector jobs increases. The combined effect is rather different for males and females. For males, formal sector preference increases even though the effect of unemployment on formal sector wages is negative. For only 2 percent of the males the preferred sector changes from working to non-working. Because the probability of rationing increases, not all of the workers end up working in the "good" formal sector. All effects considered, "good" formal and informal sector participation decrease and the participation in the other sectors increase. For females, the effect of an increase of unemployment is clearly a move out of employment towards non-participation. Both the "preferred" as well as the "actual" numbers decrease for both sectors. The simulation shows that females' labour supply is more sensitive to labour market conditions than that of men.

5.7 Concluding Remarks

Labour supply in urban areas of a developing country has been analyzed using a model that explicitly incorporates rationing for formal sector jobs. Information on search, including on-the-job search, allows us to disentangle the effects of sector preference and rationing in the model. The informal sector is taken as a free entry sector. The test of labour market segmentation that follows is thus a test for rationing of formal sector jobs, and not a test for wage differentials between sectors, as is common in the literature. The model corrects for endogenous selection into sectors and treats earnings in all sectors as endogenous. In general, estimation results are in line with the segmented labour market hypothesis.

Rationing is treated as a heterogeneous phenomenon in the model. The results show that the probability of rationing decreases with additional education. Conclusions on whether labour market segmentation is present vary with observed characteristics of the individual. Ethnicity has a significant positive effect on the probability of rationing for both males and females. The effect is the largest for females. Regional variation is

exploited to estimate demand side effects on earnings and participation. Larger labour markets increase the probability of rationing for males. Unemployment has a positive effect on rationing for both sexes.

Simulation exercises can be done to investigate the effect of a change in exogenous variables on earnings and participation in sectors. The structure of the model has the advantage that the effects resulting from changes in preferences, wages and rationing can be analyzed separately. For males, exogenous shocks mostly influence the choice between formal and informal sector participation and not so much the choice of whether to work or not. For females, on the contrary, the move is more in and out of employment. The latter is in contradiction with time series macro figures. Over the period from 1976 to 1988, over which there was a sharp decline in economic growth, the female participation rate in rural[29] areas increased from 16 to 28 percent (UDAPE 1991). (Male participation decreased slightly, in line with the results). This suggests that females' labour supply decisions are not only driven by their wage offers. Perhaps they change their labour supply in response to shocks that affect other persons within the household. They may compensate for the loss of income of the husband, for instance. These, intra-household, labour supply effects will be analyzed in the next chapter.

[29] Urban figures would have been more appropriate to report but are unavailable. Rural figures are informative because the decline in economic growth was nationwide.

Appendix 5A

Table 5.A.1. Descriptive statistics
(means and sample fractions; standard deviations in parentheses)

	male			female		
	formal	inform	not working	formal	inform	not-working
highest level of education attained:						
basic/none	0.21	0.35	0.24	0.11	0.40	0.29
inter	0.14	0.19	0.13	0.08	0.15	0.15
medio	0.29	0.30	0.33	0.22	0.20	0.29
midtech	0.04	0.03	0.04	0.09	0.03	0.05
hightech	0.03	0.02	0.03	0.04	0.01	0.01
normal	0.06	0.01	0.02	0.26	0.02	0.03
university	0.19	0.07	0.16	0.17	0.03	0.05
other	0.04	0.03	0.05	0.03	0.16	0.13
married	0.79	0.84	0.59	0.56	0.71	0.80
ethnic	0.30	0.39	0.33	0.20	0.46	0.34
age	36.1	39.5	38.7	34.0	39.3	36.8
hourly earnings[30]	2.50	2.38		2.03	1.94	
	(4.7)	(4.0)		(2.6)	(4.4)	
per cap net dissavings (/100)	0.01	-0.03	0.79	0.24	-0.05	0.37
	(2.23)	(1.78)	(1.71)	(2.77)	(1.63)	(1.99)
number of observations	3705	1763	881	1494	1917	3882

[30] primary activity

Table 5.A.2 Sector preference section

	males				females			
	formal vs non-part		informal vs non-part		formal vs non-part		informal vs non-part	
	parameter	std err	parameter	std err	parameter	std err	parameter	std err
cnst	0.961	0.242	-1.601	0.396	-0.421	0.306	-1.499	0.373
age	0.036	0.013	0.096	0.019	0.101	0.016	0.123	0.017
age squared /100	-0.075	0.016	-0.094	0.023	-0.160	0.021	-0.134	0.020
young	0.037	0.014	0.020	0.018	-0.047	0.014	-0.032	0.016
prime	-0.121	0.014	-0.142	0.019	-0.045	0.013	-0.117	0.016
old	0.034	0.063	0.183	0.086	-0.029	0.062	-0.095	0.080
married	-0.028	0.171	0.838	0.225	-1.812	0.143	-0.624	0.185
net pc dissavings /1000	-5.326	0.468	-7.169	0.699	-1.346	0.462	-4.804	0.680
pred error	4.789	0.468	5.764	0.675	1.162	0.478	4.241	0.679
married*age	0.007	0.004	-0.016	0.006	0.021	0.004	0.002	0.004
σ^2	0.617	0.044	1.306	0.201	0.821	0.066	1.255	0.121

Table 5.A.3. Wage equations

	males				females			
	formal		informal		formal		informal	
	parameter	std err	parameter	std err	parameter	std err	parameter	std err
cnst	-1.106	0.209	-0.875	0.276	-1.366	0.283	-2.548	0.356
age	0.061	0.010	0.064	0.012	0.037	0.013	0.080	0.015
age squared /100	-0.067	0.012	-0.066	0.014	-0.046	0.017	-0.086	0.018
inter	0.077	0.058	0.068	0.058	0.096	0.066	0.120	0.074
medio	0.219	0.049	0.195	0.050	0.307	0.058	-0.023	0.065
midtech	0.530	0.098	0.212	0.111	0.832	0.088	-0.166	0.113
hightech	0.635	0.134	0.192	0.140	1.078	0.138	-0.032	0.172
normal	0.372	0.098	-0.573	0.116	1.292	0.088	-0.781	0.120
university	0.879	0.058	0.104	0.071	1.467	0.080	-0.177	0.107
other	0.139	0.081	-0.327	0.095	-0.015	0.080	-0.091	0.082
missed years	0.069	0.014	0.006	0.016	0.074	0.016	-0.047	0.019
econ act /10	0.208	0.028	0.094	0.031	0.137	0.032	0.174	0.038
unemployment *10	-0.081	0.110	-1.074	0.121	-0.207	0.125	-0.657	0.147
ethnic	-0.120	0.041	-0.022	0.041	-0.048	0.047	0.184	0.051
σ^2	1.031	0.043	0.790	0.034	0.929	0.056	1.502	0.085
σ_{ir}	0.701	0.040	0.078	0.040	0.472	0.066	0.786	0.092

Table 5.A.4. τ estimate-difference between good and bad formal sector wage (see equation 5.4)

	males		females	
	parameter	std err	parameter	std err
cnst	0.241	0.175	0.163	0.350
age	0.016	0.008	-0.007	0.017
age squared /100	-0.020	0.010	0.004	0.023
inter	0.005	0.047	-0.026	0.103
medio	0.054	0.040	-0.024	0.080
midtech	0.122	0.077	0.119	0.119
hightech	0.251	0.084	-0.051	0.196
normal	-0.348	0.119	-0.368	0.144
university	0.236	0.045	0.156	0.090
other	0.053	0.079	0.142	0.117
missed years	0.022	0.011	-0.012	0.021
econ act /10	0.006	0.022	-0.044	0.041
unemployment *10	0.116	0.091	0.622	0.189
ethnic	-0.059	0.032	0.071	0.063

Table 5.A.5. Rationing equations

	males		females	
	parameter	std err	parameter	std err
cnst	-0.072	0.241	0.680	0.377
age	-0.060	0.010	-0.078	0.017
age squared/100	0.068	0.013	0.076	0.023
inter	-0.117	0.061	-0.360	0.107
medio	-0.180	0.052	-0.435	0.091
midtech	-0.213	0.102	-0.752	0.126
hightech	-0.289	0.130	-0.865	0.177
normal	-0.378	0.099	-0.689	0.109
university	-0.523	0.065	-0.748	0.111
other	-0.330	0.101	0.185	0.129
missed years	-0.047	0.015	-0.061	0.024
econ act /10	0.103	0.030	0.001	0.043
unemployment *10	0.933	0.122	1.122	0.177
ethnic	0.114	0.041	0.265	0.065

6. Household Labour Supply

6.1 Introduction

We analyze labour supply behaviour of the two partners in two- adults families. We distinguish four types of labour supplied by the family: husband's and wife's hours of work in the formal sector and in the informal sector. We present a static structural model, focusing on the relation between these four types of labour supply, their sensitivity to all four wages, and other family income.

One objective of this study is to see to what extent it is possible for a household in a developing country, to smooth consumption by reallocating labour supply of the family. Since poorer families may need more than one income to reach a level of subsistence, household income is often generated by more than one member of the household. The labour supply decisions of the individual family members are likely to be correlated. For instance, if earnings of one family member fall, other family members may increase their labour supply to compensate for the loss in family income. The combined effect on household consumption is likely to be less than the initial shock.

The point of departure is a neoclassical family labour supply model. The family is assumed to take a joint decision regarding household consumption and labour supply of its members. Its objective is to maximize utility, determined by household consumption and leisure of all family members, under a household budget restriction. This approach is used in many studies. For example, Hausman and Ruud (1984) extend the linear labour supply model to the two adults family case and apply it to US data. Kapteyn et al. (1990) apply this model to Dutch data. Ransom (1987) uses a quadratic utility function to analyze labour supply of two adults households. Newman and Gertler (1991) estimate labour supply of rural households of varying size in Peru.

The focus is on urban labour markets in developing countries and we distinguish between a formal and an informal sector. This opens another way in which households can adjust income: if wages fall in one sector, some individuals may find it advantageous to switch to the other. During a period of economic malaise with a direct negative effect on formal sector wages, the informal sector is often seen to expand. This is referred to as the buffer function of the informal sector (Todaro 1989). We incorporate this phenomenon in our model.

Our structural approach implies that a separate household utility function would have to be specified for each type of household.[31] We therefore limit ourselves to households

---

[31] Even a reduced form approach, as presented in Newman and Gertler (1991), implicitly makes arbitrary assumptions on the relations between utility functions of families of different composition.

with one prime age male and one prime age female, which we refer to as two adults households. Some information on the position of these households among all households in the sample, will be presented in section 6.3.

The organization of the chapter is as follows: in section 6.2 the model is introduced. In section 6.3 we discuss the dataset used in the estimation and provide descriptive statistics. Section 6.4 contains information on the estimation strategy and the results of the estimation are presented in section 6.5. Section 6.6 concludes.

6.2 The Model

Since the few individuals with two jobs are removed from the sample, nobody in the sample works both in the formal and in the informal sector. We assume that an individual can earn a fixed hourly wage in each sector, where the wages in the two sectors can be different. The simplest assumption then is to assume that the individual simply chooses the sector with highest hourly earnings. This however is not necessarily consistent with the data (see results page 72). Unobserved non-monetary returns, representing both preferences and costs of entry, are introduced to explain why people may choose the sector yielding the lower (monetary) earnings.

As a consequence, three sections of the model have to be distinguished. The labour supply section models the joint labour supply decision of the two spouses. Labour supply functions are derived on the basis of wage rates, nonlabour income, and individual taste shifters. The wage rate of each spouse is the maximum of the wage rates he or she can earn in the formal and the informal sector, including non-monetary returns. The second section of the model describes the wage offers in both sectors for both sexes, excluding the unobserved non-monetary returns. The third part of the model consists of the non-monetary returns equations.

Labour supply The labour supply section of the model is identical to that of Ransom (1987). A household is characterized by a quadratic direct utility function which has household consumption, including non-monetary returns, and leisure of both partners as arguments. As in chapter 5, monetary and non-monetary returns are assumed to be perfect substitutes. The family maximizes utility subject to a household budget constraint and nonnegativity conditions on hours worked:

$$\begin{aligned} \max \quad & U(Z) = \alpha Z - \frac{1}{2} Z' \beta Z \\ s.t. \quad & w_m^* h_m + w_f^* h_f + Y = C^* \\ & h_m \geq 0 \\ & h_f \geq 0 \end{aligned} \tag{6.1}$$

with

$Z =$ $[T-h_m, T-h_f, C^*]'$
$T =$ time endowment
$h_m, h_f =$ hours worked by male and female
$C^* =$ family consumption, including non-monetary returns
$Y =$ nonlabour income
$w_m^*, w_f^* =$ hourly wage of male and female, including non-monetary returns
$\alpha \in \mathbb{R}^3$; $\beta=(\beta_{ij}) \in \mathbb{R}^{3x3}$.

In our case, wage rates are the maximum of formal and informal sector wages, including non-monetary returns (see below). We assume that the budget constraint is binding, i.e. that utility increases with C^*. If neither of the two nonnegativity conditions on hours worked are binding, first order conditions (marginal utility with respect to hours worked set to zero) can be written as

$$\alpha_1^*+\alpha_3^* w_m^* - \beta_{11} h_m - \beta_{33} w_m^* (w_m^* h_m + w_f^* h_f + Y) - \beta_{12} h_f + \beta_{13}(2 w_m^* h_m + w_f^* h_f + Y) + \beta_{23} w_m^* h_f = 0 \qquad (6.2)$$

$$\alpha_2^*+\alpha_3^* w_f^* - \beta_{22} h_f - \beta_{33} w_f^* (w_m^* h_m + w_f^* h_f + Y) - \beta_{12} h_m + \beta_{23}(2 w_f^* h_f + w_m^* h_m + Y) + \beta_{13} w_f^* h_m = 0 \qquad (6.3)$$

The α^*s are functions of α, β and the time endowment. See Ransom (1987) for details. The quadratic specification implies that it is not necessary to specify the time endowment. Following Ransom, we allow α_1^* and α_2^* to be a function of observed taste shifters Q_{ki} and unobserved taste shifters ϵ_{ki} ($k=1,2$), where the subscript i denotes the household:

$$\alpha_{ki}^* = Q_{ki}\Gamma_k + \epsilon_{ki} \qquad k=1,2 \qquad \epsilon_i = \begin{pmatrix} \epsilon_{1i} \\ \epsilon_{2i} \end{pmatrix} \qquad \epsilon_i \sim N(0,\Sigma) \qquad (6.4)$$

If one of nonnegativity constraints becomes binding and one of the partners does not work, one of the first order conditions changes into an inequality condition. This results in a simultaneous model of two tobit equations. Due to the quadratic utility function, the underlying latent model is linear. See Ransom (1987) for details.

Wages The log of the hourly wage in both sectors is modelled as a linear function of exogenous variables and an error term:

$$\ln(w_{kj}) = X_k \tau_{kj} + \eta_{kj} \qquad \eta_{kj} \sim N(0,\sigma_{kj}^2) \qquad \begin{matrix} j=1,2 \text{ for formal and informal} \\ k=1,2 \text{ for male and female} \end{matrix} \qquad (6.5)$$

For ease of notation we have dropped the index i as we shall continue to do in the rest of this chapter. We use separate wage equations to describe earnings in both sectors for

both sexes. The error terms are assumed to be independent of each other and of the random preference terms in the labour supply section of the model.

Non-monetary returns Non-monetary returns reflect the sum of the effects resulting from sector specific preferences and rationing. Sector specific preferences include non-monetary (and non-measured) benefits from working such as health insurance or status attached to participation in a specific sector. Following Magnac (1991) rationing is modelled as waiting queues, resulting in costs proportional to wages. Magnac assumes homogeneous preferences and interprets non-monetary returns as evidence of labour market segmentation. In this model, non-monetary returns are included in the model in exactly the same way as in Magnac, but we changed the interpretation.

We model non-monetary returns as a fraction of the monetary wage. We normalize informal non-monetary returns to zero. The log of the (formal sector) non-monetary returns NMR_k ($k=1$ (male), 2 (female)) are written as a function of individual characteristics, local labour market conditions and an error term:

$$\ln (NMR_k+1)=V_k\gamma_k+\mu_k \qquad \mu_k \sim N(0,\sigma_k^2) \tag{6.6}$$

We assume that each individual chooses the sector with highest wage offer, including non-monetary returns. For the formal sector the wage is $w_{Ik}^*=(NMR_k+1)w_{Ik}$. For the informal sector, the wage offer itself enters because non-monetary returns equal zero. With probability one, the wage offers will be different, so that the individual chooses one sector only.

The complete model thus consists of eight equations: two wage equations for each sex, one non-monetary returns equation for each sex, and two labour supply equations. Note that the labour supply section of the model is modelled in terms of hourly wages, while in the second section of the model wages and non-monetary returns are modelled in logs, so that the model as a whole is nonlinear.

Apart from modelling labour supply and intra-household effects, the model differs from the one presented in the previous chapter in a number of ways. First of all, the interpretation of non-monetary returns (NMR) is different. Non-monetary returns now reflect the combined effect of sector specific preferences and rationing. We do not distinguish between the two in the estimation and thus ignore information on search in the dataset. Second, the unobserved reservation wages are now defined in the model by the first order conditions (6.2) and (6.3). Hence, only the difference in non-monetary returns between sectors is identified. Third, we use other income instead of net dissavings. For coherency reasons we would have preferred to use net dissaving. However, we were unable to obtain estimates this way (the maximization routine did not converge to a maximum).

The model is an improvement on the Ransom (1987) model in that the wage equations are incorporated and estimated jointly with the labour supply equations. Ransom predicted wage offers for nonparticipants using a separate model, thus ignoring wage rate prediction errors. Furthermore, we distinguish between two sectors. The model also generalizes the one presented in Magnac (1991): we follow a more structural approach, and consider not only participation but also hours worked. Moreover, while Magnac considers individuals, we work on the household level and analyze intra-household interactions.

6.3 Data

The EIH89 survey covers 7264 households in 8 urban centers (cf chapter 2). 3712 households contained one prime age male and one prime age female, with both of them potential workers (between 19 and 65 years old, in good health and not attending full-time education). The definition of formal and informal sector is the same as in the previous chapter (see page 18 and 67), ie with an asset threshold for self-employed workers. 476 households were excluded because one of the partners could not be classified, 281 because one of the partners had more than one job. Finally, 153 households were excluded due to missing information on one of the necessary variables (wages (123 cases), education level, hours worked).[32] The sample used for estimation thus contains 2802 families. Table 6.1 contains some descriptive statistics. 91 percent of all males is working, against 40 percent of all females. The majority of males work in the formal sector. The opposite holds for females. Most individuals participating in the informal sector have little education; the higher educated are overrepresented in the formal sector.

Table 6.2 contains some *prima facie* evidence of intra-household effects of labour supply. We have calculated the average female participation rate and the average log wage of working females by wage quintile of the male. The female participation rate is highest in the lower and upper male wage brackets. Yet, female wages increase with the male's wage. The high participation rate in the lower quintile could be explained by the low income earned by the male: one income is not enough to support the family. For the high male wage bracket, the own wage effect of the female seems to dominate: the high wage she can earn induces her to work.

---

[32] For the sample of 3712 households consisting of one prime age female and one prime age male, average per capita household consumption is 142.8. For the selected sample, it is 136.0. This suggests that the data cleaning does not lead to serious sample selectivity problems. Compared to the average in the total survey, per capita income of two adults households is about 8 percent lower.

Table 6.1 Descriptive statistics
(formal and informal sector definition based on worker's status, means and sample fractions; standard deviations in parentheses)

	male			female		
	formal	inform	not working	formal	inform	not working
highest level of education attended:						
basic	0.22	0.36	0.32	0.10	0.44	0.35
inter	0.15	0.21	0.13	0.09	0.16	0.16
medio	0.30	0.28	0.30	0.20	0.20	0.28
middle technical	0.04	0.02	0.03	0.08	0.03	0.04
higher tech	0.02	0.02	0.03	0.04	0.01	0.01
normal	0.06	0.02	0.01	0.31	0.01	0.02
university	0.16	0.06	0.15	0.15	0.02	0.04
other	0.05	0.03	0.04	0.04	0.12	0.10
married	0.96	0.98	0.92	0.96	0.94	0.97
ethnic	0.33	0.42	0.36	0.22	0.48	0.34
age	35.66	37.63	39.58	34.33	35.55	33.06
wage	2.21	2.16		1.90	1.72	
	(3.16)	(2.55)		(1.54)	(2.29)	
hours	52.1	53.6		37.4	45.7	
	(17.1)	(17.8)		(16.2)	(25.7)	
other income (/100)	7.46	4.82	35.66	9.01	5.97	10.94
observations	1755	782	265	423	710	1669

Table 6.2 Female labour supply by male wage quintile (participation rate in percentages, log female's wage conditional upon working, standard errors in parentheses)

wage quintile male	participation rate female	log wage female
not working	49.4	0.13
	(3.5)	(0.07)
1 (poor)	44.3	-0.34
	(2.2)	(0.06)
2	36.9	-0.02
	(2.1)	(0.06)
3	39.1	0.26
	(2.2)	(0.06)
4	36.7	0.32
	(2.1)	(0.06)
5 (rich)	40.6	0.78
	(2.2)	(0.06)

6.4 Estimation

Due to its nonlinear nature, the model cannot be estimated by maximum likelihood. Exact likelihood contributions would be too hard to compute at each step of the

maximization process. Instead, we maximize an approximation of the likelihood, based upon simulations of some of the errors in the wage and non-monetary returns equations. This method is an example of smooth simulated maximum likelihood (cf. Boersch-Supan and Hajivassiliou (1993), van Praag and Hop (1987), for example). We only describe the main idea here. Details can be found in the appendix.

If, for a given family, both partners are working and w_m^* and w_f^* are known, the likelihood contribution of the labour supply part of the model (conditional on wages) is identical to that in Ransom (1987). We denote it by $L_I(h_m, h_f; w_m^*, w_f^*)$. To keep notation simple, we suppress the other arguments it depends on (taste shifters, other family income). The complete likelihood contribution is then given by

$$L(h_m,h_f,s_m,s_f,w_m^*,w_f^*) = L_I(h_m,h_f|w_m^*,w_f^*)L_{II}(s_m,w_m^*)L_{III}(s_f,w_f^*) \tag{6.7}$$

Here L_{II} and L_{III} are the likelihood contributions of the wage and non-monetary returns equations for male and female,respectively. Since we do not include any correlation between the wages and non-monetary returns of males and females in the model, L_{II} and L_{III} can be written as a product. s_m and s_f are the observed sectors of male and female. Because of the linearity of this part of the model and independence assumptions of the errors, L_{II} and L_{III} can easily be computed. It factors out into a contribution of the male and a contribution of the female, both similar to Magnac's (1991) likelihood. In this case therefore, the likelihood contribution can be written down exactly.

This is only relevant however, if both partners work in the informal sector. If either of them does not participate or works in the formal sector, his or her relevant wage is not observed. In case of nonparticipation, this is the familiar problem in estimating structural labour supply models with unobserved wage rates (cf. MaCurdy et al (1990), for example). In case of formal sector work, we observe the monetary wage, but we must include non-monetary returns which are always unobserved. Let us assume that the male and the female both work in the formal sector. The likelihood contribution can now be written as

$$L(h_m,h_f,s_m,s_f,w_m,w_f) = E\{L_I(h_m,h_f|w_m^*,w_f^*)L_{II}(s_m,w_m^*)L_{III}(s_f,w_f^*)\} \tag{6.8}$$

where the expectation is taken with respect to the unobserved errors in the non-monetary returns equations, linking the observed w_m and w_f to the unobserved w_m^* and w_f^*. See appendix 6A for details. This expectation cannot be computed analytically, because L_I is a complicated nonlinear function of w_m^* and w_f^*. It is therefore replaced by a simulated mean:

$$L_H(h_m,h_f,s_m,s_f,w_m,w_f) = \frac{1}{H}\sum_{j=1}^{H}\{L_I(h_m,h_f|w_{mj}^*,w_{fj}^*)L_{II}(s_m,w_{mj}^*)L_{III}(s_f,w_{fj}^*)\} \tag{6.9}$$

where w^*_{mj} and w^*_{fj} ($j=1,\ldots,H$) result from adding H independent draws of the errors in the NMR equations to the observed formal sector wages.

This procedure can easily be generalized to other cases. For those individuals who do not participate, we draw all the errors in the wage and non-monetary returns equations. The maximum wage offer from the two sectors is substituted in the labour supply part of the model. This type of nonlinearity was also solved through simulation by Laroque and Salanie (1989) for a disequilibrium model. The sample log likelihood is replaced by an approximate log likelihood, replacing each likelihood contribution by its simulated approximation. The simulated maximum likelihood (SML) estimator maximizes the approximate sample loglikelihood.

The estimator is consistent if both the number of observations N and the number of draws per observation H go to infinity. Moreover, if draws for different observations are independent, the estimator is asymptotically efficient if $\sqrt{N}/H$ goes to zero (Gourieroux and Monfort 1993). Because the errors in the labour supply part of the model are retained and not simulated, the approximate likelihood is a continuous and differentiable function of the parameters. This makes maximization feasible and, according to previous studies on similar models (Boersch-Supan and Hajivassiliou, 1993, for example), should lead to acceptable results for small values of H already. Just like in the previous chapter we use $H=30$.

6.5 Estimation Results

We first present the estimates of the labour supply model, equations (6.1) through (6.4). The estimates of Γ_1 and Γ_2, determining α^*_1 and α^*_2, are in table 6.3. A higher value of α^* is associated with a higher propensity to work, since α^* is a negative function of α. Exogenous variables included in Q_k ($k=1,2$) refer to family composition and age. The number of children in the household (YOUNG) significantly increases the propensity to work for males and has the opposite effect for females. The number of older persons in the household (OLD) is insignificant. A quadratic age pattern is significant for both males and females.[33] The propensity to work is highest at age 33 for males and at age 42 for females. There is a significant negative correlation between the two random preference terms indicating, *ceteris paribus*, a tendency towards specialization.

[33] α^*_1 and α^*_2 depend on the male's and female's age, respectively.

Table 6.3 Estimates for taste shifters in labour supply model (standard errors in parentheses)

	α_1^*	α_2^*	
cnst	1.546	-8.652	
	(0.733)	(1.830)	
young	0.125	-0.154	
	(0.037)	(0.052)	
old	-0.067	0.196	
	(0.306)	(0.318)	
age	1.419	3.843	
	(0.390)	(0.935)	
age squared/100	-2.173	-4.592	
	(0.471)	(1.145)	
σ^2	9.117	7.068	σ_{12} (Σ_{12}) -1.944
	(0.416)	(2.077)	(0.802)

The estimates for the matrix β are presented in table 6.4. To normalize the scale of the utility function, we have set β_{11} equal to one, following Ransom (1987). All estimated coefficients are significant. The marginal utility of leisure increases with additional leisure of the partner. Moreover, marginal utility of leisure of both partners increases with family consumption.

Table 6.4 Estimates for β matrix (standard errors in parentheses)

	male leisure	female leisure	consumption (/10)
male leisure	1		
female leisure	-0.262	0.414	
	(0.054)	(0.088)	
consumption(/10)	-0.340	-0.326	0.0178
	(0.057)	(0.055)	(0.0029)

In writing down the first order conditions (6.2)-(6.3), we do not take account of the possibility that an interior point of the budget set is chosen. It is thus assumed that the utility function increases with family consumption. According to our results, this is the case for 2757 out of 2802 observations. The 45 remaining observations are discarded in the simulations below. If utility increases with family consumption, the solution of the Lagrange equations (equalities and inequalities) corresponds to the utility maximum if the utility function is quasiconcave. Positive definiteness of β is a sufficient but not a necessary condition. Since our estimate of β is not positive definite, we checked concavity conditions for all observations. It appears that they are satisfied without exception.[34]

[34] This also implies that the model is coherent, in the sense that endogenous variables are uniquely determined (cf. Van Soest et al., 1993).

Figures 6.1 and 6.2 illustrate the shape of the labour supply curves. Figure 6.1 shows unconditional supply curves, i.e. predicted numbers of hours worked (divided by 10) for males and females as a function of both partners' wages (including non-monetary returns).[35] α_1^* and α_2^* are set equal to their sample means (α_1^*=4.1, α_2^*=-1.7), random preferences are set equal to zero. Male labour supply is forward bending in most of the range of male wages, and backward bending for very high wage rates, where the income effect dominates the substitution effect. If w_f^* is low, the female does not work, and male hours depend on w_m^* only. If the female works, male hours are negatively affected by the female's wage. The female's hours of work increase with her own wage and decrease with the husband's wage.

Figure 6.2 shows the probability of participation for both males and females as a function of their own wage and the number of hours worked by the partner (divided by 10). The wage of the partner and all family characteristics are kept at the mean predicted value. Random preferences are taken into account. For females, the own wage effect on the probability of working is positive. The effect is stronger if the male works fewer hours. The effect of the husband's wage on the wife's hours is small, and its sign depends on the female's wage. Only for high wages of the female, this effect is substantially negative. For the family the graph refers to, the probability that the male works is always higher than 0.95. There is a slight positive own wage effect and a negligible effect of the female's hours on this probability.

The estimates of the wage equations and the non-monetary return equation for both sexes are presented in table 6.5. In the wage equations we have included individual characteristics such as age, education level and ethnicity, and variables describing local labour market conditions such as the local unemployment rate and a measure for the size of the economically active population, as a proxy for the size of the local labour market. The specification is identical to the previous chapters and the results are similar. For example, returns to education are larger in the formal sector than in the informal sector. This may indicate that the formal sector requires skills obtained through the formal education system or that education is used as a screening device in the formal sector. A larger local labour market generally leads to a higher wage. The effect is significant for males in the formal sector and for females in the informal sector. The significantly negative effect of the local unemployment rate is largest in the informal sector. This can be explained by the fact that the informal sector is more competitive than the formal sector. In the formal sector, ethnic minorities are paid significantly less than others.

In the non-monetary return equations for formal sector employment, we have included the variables of the wage equations and two taste shifters: the numbers of young and old persons in the family. Non-monetary returns may result from demand side constraints (rationing) or from individual preferences concerning sector participation. For both male

[35] Estimated sample averages of wages including nonmonetary returns are 6.6 for males and 3.0 for females.

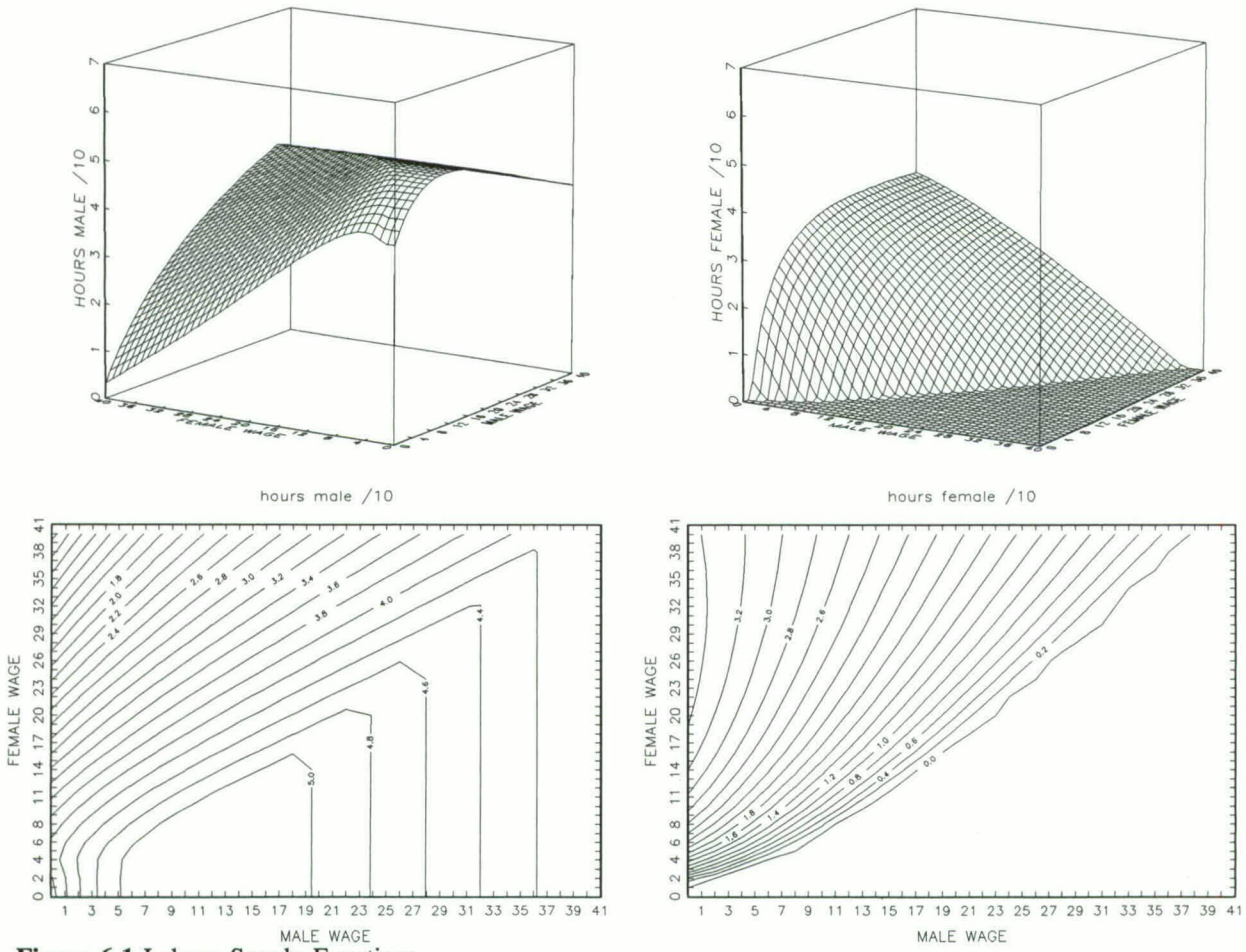

Figure 6.1 Labour Supply Functions.

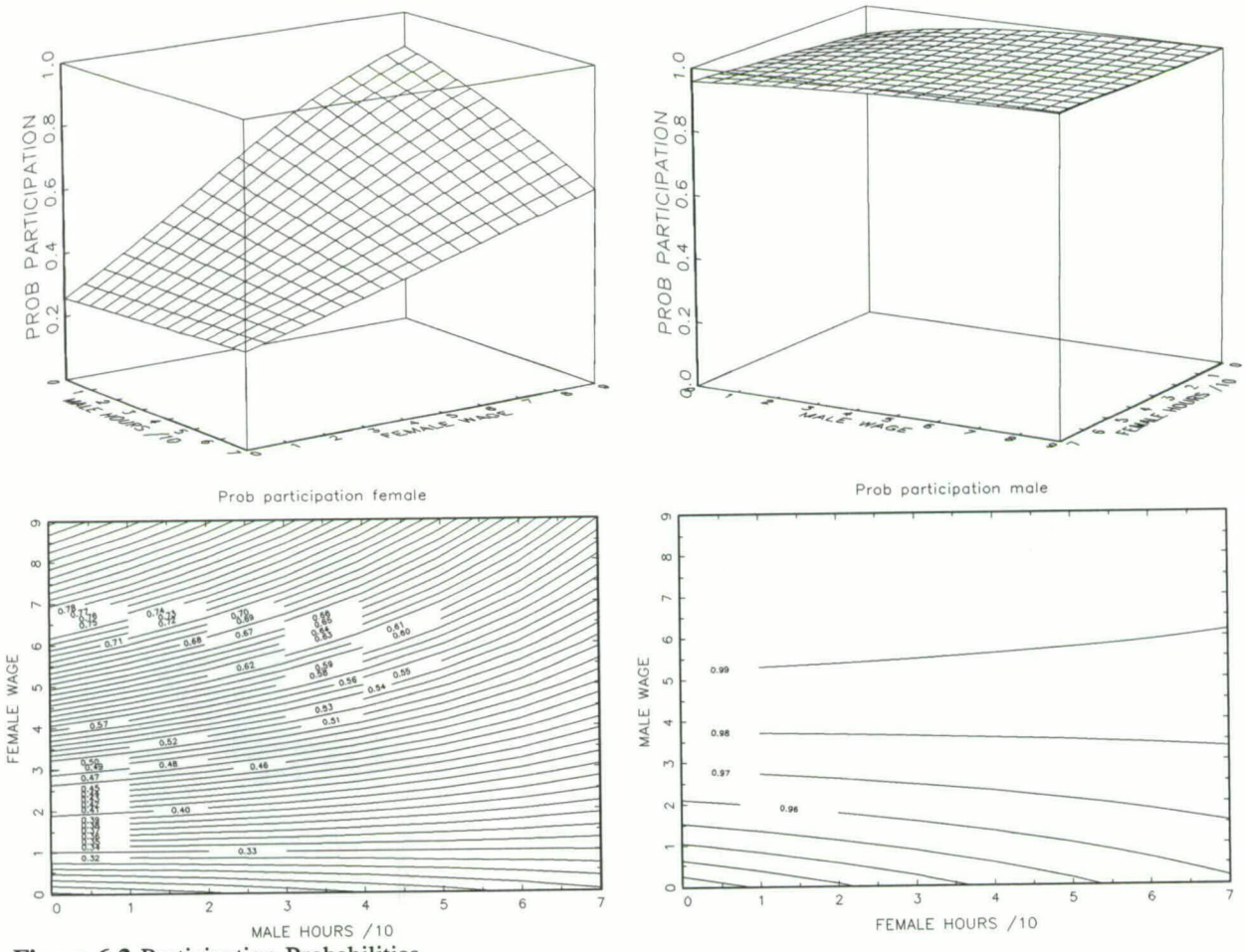

Figure 6.2 Participation Probabilities.

Table 6.5 Estimated coefficients for wage equations and non-monetary return equations (standard errors in parentheses)

	male			female		
	formal	informal	nmr	formal	informal	nmr
cnst	-1.334	0.649	2.399	-2.069	-0.820	-2.055
	(0.281)	(0.384)	(0.372)	(0.649)	(0.514)	(0.944)
age	0.710	0.101	-0.124	0.808	0.429	1.109
	(0.137)	(0.168)	(0.206)	(0.318)	(0.239)	(0.502)
age squared	-0.771	-0.083	-0.259	-0.967	-0.483	-1.441
	(0.177)	(0.202)	(0.269)	(0.410)	(0.299)	(0.647)
inter	0.167	-0.059	-0.241	0.553	0.177	0.365
	(0.060)	(0.076)	(0.089)	(0.129)	(0.103)	(0.202)
medio	0.351	0.168	0.182	0.812	0.210	0.395
	(0.051)	(0.070)	(0.072)	(0.119)	(0.091)	(0.192)
midtech	0.664	0.206	-0.049	1.120	0.172	1.000
	(0.110)	(0.172)	(0.196)	(0.142)	(0.204)	(0.217)
hightech	0.784	0.311	-0.046	1.165	0.688	1.448
	(0.126)	(0.223)	(0.209)	(0.225)	(0.264)	(0.264)
normal	0.581	-0.375	0.273	1.280	-0.068	2.311
	(0.105)	(0.175)	(0.131)	(0.126)	(0.254)	(0.182)
university	1.181	0.099	-0.159	1.767	0.326	1.130
	(0.062)	(0.102)	(0.094)	(0.134)	(0.177)	(0.208)
other	0.583	-0.190	0.347	0.016	-0.253	-0.377
	(0.091)	(0.170)	(0.098)	(0.169)	(0.124)	(0.283)
miss years	0.068	0.027	-0.001	0.123	0.013	0.121
	(0.015)	(0.022)	(0.021)	(0.031)	(0.027)	(0.049)
econ act	0.144	0.041	-0.226	0.089	0.199	-0.010
	(0.029)	(0.044)	(0.044)	(0.055)	(0.054)	(0.075)
unempl	-0.630	-1.211	-0.351	-0.568	-0.918	-0.507
	(0.121)	(0.173)	(0.162)	(0.226)	(0.219)	(0.308)
ethnic	-0.160	-0.031	-0.220	-0.210	0.041	-0.291
	(0.044)	(0.060)	(0.060)	(0.091)	(0.078)	(0.134)
young			-0.059			-0.088
			(0.017)			(0.037)
old			-0.083			-0.142
			(0.080)			(0.136)
σ	0.751	0.798	1.086	0.684	0.904	1.113
	(0.010)	(0.020)	(0.025)	(0.025)	(0.023)	(0.054)

and females we find that "normal" education (primary school teachers) has positive non-monetary returns. People with this type of training attach a higher status to teaching in a primary school, which is exclusively formal, to informal sector work. If one assumes that preferences do not depend on ethnicity, the negative coefficient on "ethnic" can be interpreted as an indicator of relatively high job search costs for ethnic minorities for formal sector jobs or, discrimination. The number of young children increases preference for informal sector jobs. Perhaps this is because of higher flexibility of when and how much to work in the informal sector. Maybe also, additional income can be generated in the informal sector by child labour. The hypothesis of weakly competitive markets (all parameters in non-monetary returns equation equal zero) as defined by Magnac (1991), is clearly rejected for both males and females.[36]

Simulations To see how well the model predicts the distribution of hours worked and sector participation, we present figure 6.3 and table 6.6. For all observations in the sample we have simulated wages, non-monetary returns, and the number of hours worked, taking one draw from the distribution of the error terms. In the top and bottom panel of figure 6.3, the predicted and actual sample distribution of hours worked are shown, respectively. Note that the scale of the vertical axis varies per graph. Actual hours distributions for males are peaked at 40 and 48 hours per week. These peaks are not fully reproduced by the predictions. This would require a model incorporating restrictions on hours worked, as in Dickens and Lundberg (1993). For females, the distribution of both actual and predicted hours worked is more dispersed.

In table 6.6 we present a cross tabulation of the sector in which husband and wife participate.[37] We compare actual and predicted numbers. The model underpredicts the number of nonparticipants. This is particularly serious for males. Possible explanations may be fixed costs of working or a lack of available part-time jobs. The ratios of formal and informal sector participation rates are predicted with reasonable accuracy.

Table 6.6 cross tabulation of sector participation for male and female (actual and predicted, percentages)

female⇒	formal		informal		not-working		total	
male⇓	actual	pred	actual	pred	actual	pred	actual	pred
formal	11.1	10.0	12.0	17.0	39.7	39.9	62.8	66.9
informal	2.7	4.8	9.0	9.5	16.0	15.3	27.6	29.7
not working	1.4	1.3	3.4	0.8	4.9	1.4	9.6	3.5
total	15.1	16.1	24.4	27.3	60.5	56.7	100	100

[36] Wald tests statistics are 230 and 580, exceeding critical values of χ^2_{15}, for any sensible significance level.

[37] Tables 6.6 through 6.8 are based on the 2757 observations for which utility increases with family consumption.

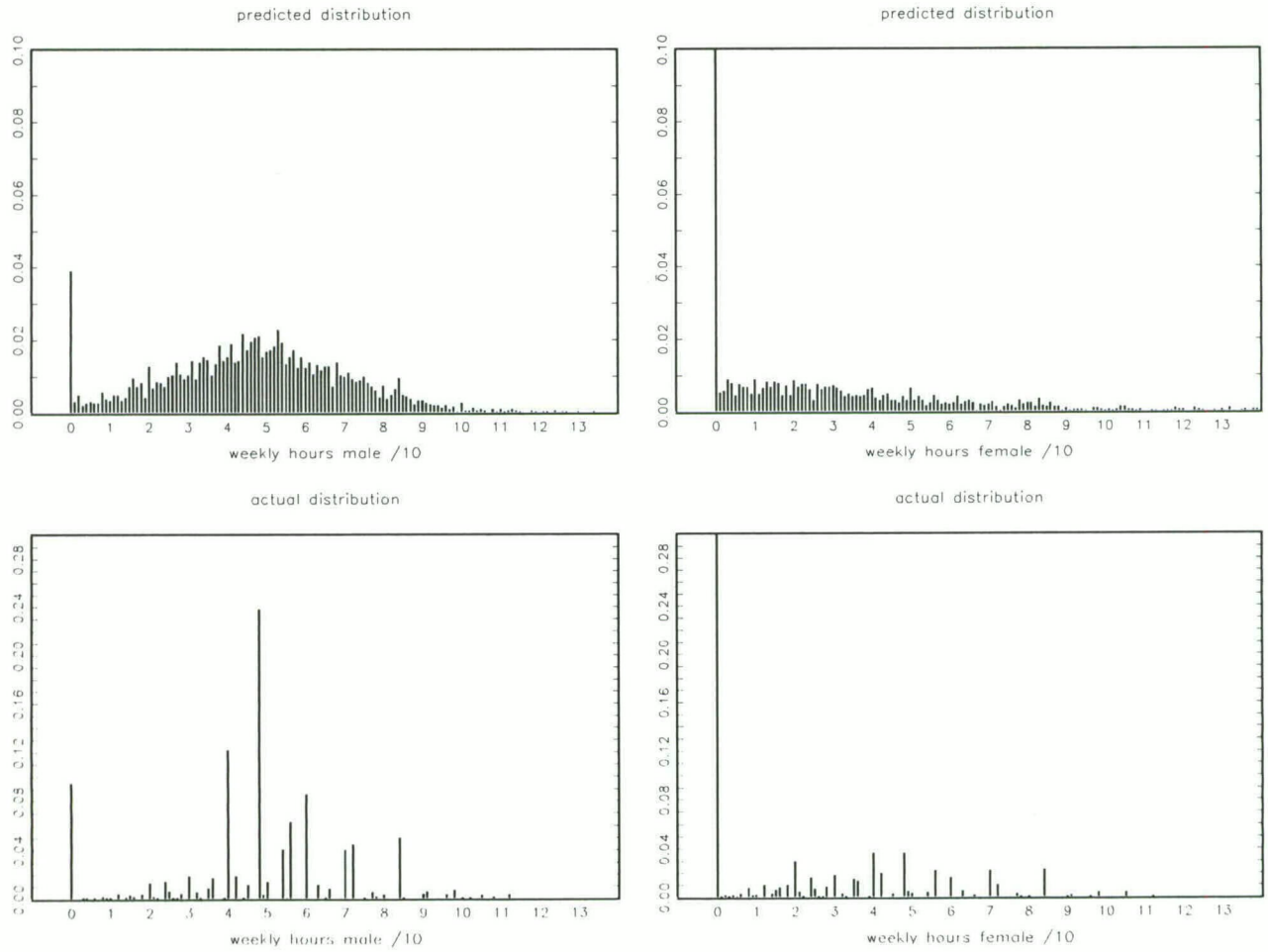

Figure 6.3 Hours Distributions.

In Table 6.7 we present the results of some simulation exercises. The objective of the first two simulations is to examine the importance of intra-household effects. We first simulate the labour supply response of a 10 percent drop of wage rates for all males. This has hardly any effect on the average number of hours the male works. Participation of males slightly decreases. Labour supply of the females however, shows a stronger response: the average number of hours worked increase by 3.0 percent, corresponding to a cross labour supply elasticity of -0.3.

A closer look at the own labour supply response for males reveals that the low elasticity is not uniform over the sample. Males with a positive labour supply response are those who initially had a high wage. For most males with a low wage, the labour supply response to a wage decrease is negative. This corresponds with the inverted U shape of the labour supply curve shown in figure 6.2.

The second simulation concerns a 10 percent decrease of wage rates of females. This has only a very small effect on hours worked by males and females. Male hours increase and female hours decrease, but both effects are less than one percent. To get some insight in aggregate income elasticities, we also performed a simulation in which nonlabour income increased by 10 percent for all households. For 15 percent of the households this does not have any effect, since their non-labour income was zero to start with. The effects were quite small. For both males and females, hours worked decrease slightly. Income elasticities are 0.005 for males and 0.012 for females.

A main objective of this chapter is to see how well a family can counter an exogenous shock in income by adjusting hours worked and by switching between sectors. Therefore, we simulated a fall in all formal sector wages by 10 percent. See tables 6.7 and 6.8. There is a strong direct effect on the income of the majority of males employed in the formal sector. 2.3 percent of males switch from the formal to the informal sector, while the participation rate for males remains the same. Labour force participation of females increases by 0.4 percent. This is a result of two opposite effects: an increase due to reduced partner's earnings (cf. previous simulation), and a decrease due to the fall of the own (formal sector) wage. The average number of hours worked increases for both males and females. Labour supply of males increases because formal sector workers are generally in the higher income brackets. The size of the informal sector increases by 5.9 percent. This prediction should be viewed as an upper limit, since a fall in formal sector wage offers will often be accompanied by a fall in informal sector wages due to a lower demand for services from the informal sector.

Table 6.7 Simulations (sample averages and changes of sample averages, standard errors in parentheses[38], wages and incomes excluding non-monetary returns, Δ = difference from predicted value)

	actual	predicted	wage male -10%	wage female -10%	formal wage -10%
		mean	Δ	Δ	Δ
wage offer male		2.05	-0.205	0	-0.135
		(0.04)	(0.004)		(0.005)
wage offer female		1.48	0	-0.148	-0.051
		(0.04)		(0.004)	(0.004)
hours male	47.10	46.93	-0.034	0.156	0.120
		(0.55)	(0.040)	(0.015)	(0.031)
perc working	90.40	96.44	-0.208	0.144	-0.009
		(0.30)	(0.081)	(0.072)	(0.076)
hours female	16.10	16.86	0.458	-0.129	0.306
		(0.77)	(0.043)	(0.032)	(0.035)
perc working	39.50	43.33	0.902	-0.838	0.419
		(1.03)	(0.191)	(0.177)	(0.171)
household income	126.30	130.95	-8.581	-2.688	-6.285
		(2.23)	(0.227)	(0.152)	(0.278)

Table 6.8 Simulated Sector participation rates after 10 percent drop formal sector wage (percentages, standard errors in parentheses and changes in percentage points (Δ))

	male		female	
	before	after	before	after
	mean	Δ	mean	Δ
formal	66.8	-2.23	16.0	-0.74
	(0.95)	(0.25)	(0.80)	(0.18)
informal	29.6	2.22	27.3	1.15
	(0.93)	(0.25)	(1.07)	(0.19)
not working	3.6	0.01	56.7	-0.42
	(0.29)	(0.08)	(1.03)	(0.17)

6.6 Conclusions

We have analyzed labour supply behaviour and the choice between formal and informal sector work of the two spouses in two adults families in urban areas of Bolivia. For this purpose, we have developed a static neoclassical model, combining the family labour supply model of Ransom (1987) with the model explaining sector choice and earnings of Magnac (1991). Our main empirical conclusion is that intra-household effects are substantial: low earnings of the husband are compensated by more working hours of the wife. Second, we find that a wage shock in the formal sector would induces people to

[38] Based on 300 draws from estimated asymptotic distribution of estimator of β.

move to the informal sector, while the effect on nonparticipation is much smaller. Third, we find that non-monetary returns in the formal sector are usually positive. This implies that, on average, if formal and informal sector wage are equal, people prefer a formal sector job. It can be explained by differences in job characteristics. This finding is different from that of Magnac (1991), who finds that non-monetary returns are insignificant.

Although our model captures some features of the data quite well, a simulation makes clear that it is not fully capable to reproduce the data. In particular, nonparticipation of males and, to a lesser extent, of females is underpredicted. Allowing for fixed costs of working or taking account of constraints on hours worked, might help to overcome this problem. The quadratic specification of the utility function, together with the estimation method of smooth simulated maximum likelihood, make these extensions feasible areas of future research.

Appendix 6A. Simulated Likelihood Contributions

The likelihood contributions consist of three parts, L_I, L_{II}, and L_{III}, as introduced in (6.7). L_I is the contribution of the labour supply section of the model, for given wages w_m^* and w_f^* (including *NMR*) of males and females. The expression for L_I is given in Ransom (1987). L_{II} and L_{III} reflect the likelihood contributions of wage equations and non-monetary returns, for males and females, respectively. L_{II} and L_{III} are stated in terms of log wages and non-monetary returns. The jacobian of the transformation of the density functions of $\log(w^*)$ to w^* does not depend of the parameters of the model can be ignored in the formulation of the likelihood contribution. Because males and females are treated identically in the model, L_{II} and L_{III} are similar. We first consider the male and look at L_{II}.

If the male works in the formal sector, w_m is observed, but w_m^* is not, because of the non-monetary returns. If μ_1, the error in the NMR equation, were known, the likelihood contribution of this section of the model would be given by

$$L_{II}(s_m, w_m^*(\mu_1)) = \int_{-\infty}^{\ln(w_{11})+V_1\gamma_1+\mu_1-X_1\tau_{12}} f(\eta_{12})d\eta_{12}\, f(\eta_{11}) \tag{6A.1}$$

where f denotes the (normal) p.d.f. of η_{12} and η_{11} respectively. Given τ_{11}, η_{11} is observed and equals $\ln(w_{11})$-$X_1\tau_{11}$. The integral corresponds to the condition that the informal sector wage offer is less than the formal sector wage offer, including non-monetary returns. L_{II} is thus easy to compute.

If the male works in the informal sector, we observe w_m^*. Non-monetary returns in the informal sector are zero. The likelihood contribution equals

$$L_{II}(s_m, w_m^*) = \int_{-\infty}^{\ln(w_{12})-X_1\tau_{11}-V_1\gamma_1} f(\mu_1+\eta_{11})\; d(\mu_1+\eta_{11})\; f(\eta_{12}) \tag{6A.2}$$

with $\eta_{12}=\ln(w_{12})-X_1\tau_{12}$, which is observed given τ_{12}. The integral corresponds to the condition that the formal sector wage offer, including non-monetary returns, is less than in the informal sector. If the male does not participate, we do not know whether the informal or the formal sector wage is relevant, and we must condition on η_{11}, η_{12} and μ_1. L_{II} equals 1 and vanishes. The wage that enters into L_I equals

$$w_m^*(\eta_{11},\eta_{12},\mu_1) = \exp(\max(X_1\tau_{11}+\eta_{11}+V_1\gamma_1+\mu_1, X_1\tau_{12}+\eta_{12})) \tag{6A.3}$$

L_{III} is calculated in a similar way. The full likelihood contribution of the family is given by the expectation of the product of L_I, L_{II} and L_{III}, with respect to the error terms that

we conditioned on. For example, if the husband works in the formal sector and the wife does not participate, the exact likelihood contribution is given by

$$L(h_m,h_f,s_m,w_m) = \int L_I(h_m,h_f|w_m^*(\mu_1),w_f^*(\eta_{21},\eta_{22},\mu_2))L_{II}(s_m,w_m^*(\mu_1))\, f(\mu_1,\eta_{21},\eta_{22},\mu_2)\, d\mu_1 d\eta_{21} d\eta_{22} d\mu_2 \tag{6A.4}$$

where f denotes the (normal) p.d.f of $(\mu_1,\eta_{11},\eta_{22},\mu_2)$. The four dimensional integral in (6A.4) cannot be computed analytically, because L_I is a complicated nonlinear function of w_m^* and w_f^*. It is therefore replaced by a simulated mean:

$$L_H(h_m,h_f|s_m,w_m) = \frac{1}{H}\sum_{j=1}^{H} \{L_i(h_m,h_f|w_m^*(\mu_{1j}),w_f^*(\eta_{21j},\eta_{22j},\mu_{2j}))L_{II}(s_m,w_m^*(\mu_{1j}))\} \tag{6A.5}$$

where $(\mu_{1j},\eta_{21j},\eta_{22j},\mu_{2j})$, $j=1,\ldots,H$, are i.i.d. draws from the distribution of $(\mu_1,\eta_{21},\eta_{22},\mu_2)$. Other cases are treated in a similar way. The integral to be replaced varies from six dimensional ($h_m=h_f=0$) to zero dimensional (male and female work in the informal sector).

7. Conclusions

7.1 Introduction

What have we learned about the functioning of the labour market in Bolivia? Chapters 4,5 and 6 have highlighted various aspects of labour supply behaviour in urban Bolivia. Each chapter had its own focus. Chapter 4 contained a reduced form analysis of earnings in both the formal and informal sector. In chapter 5 we developed a more structural model to test for labour market segmentation. The model disentangled the effects working through individual preferences and barriers to entry for the formal sector. In chapter 6 the focus was on intra-household effects. The chapters have a lot of similarities. They are all on the basis of the same dataset. The applied models have some overlap as well. They use for example the same earnings functions. The results are not always easy to compare. In this chapter we summarize the results, taking into consideration the issues that were introduced in chapter 1.

7.2 Returns to Education

Estimates for the returns to education can be obtained from the estimates of the hourly earnings equations (Mincer 1974). The specification of the earnings equation allows us to compare the rates of return for different educational careers. Chapters 4,5 and 6 present comparable, but different estimation results for the earnings equation. In order to obtain *one* estimate for the returns to education for different schooling types we combine the results of chapter 4 and 5. We do not use the results in chapter 6 as they are based on a subsample of 2 adult households. Table 7.1 presents the average private returns to one year of education after primary (basic) education, i.e. they are averaged over the years needed to complete the relevant education level. The numbers are calculated assuming that the individual has completed the training. If the individual did not complete the training, the negative return for one missed year is given in the last row. Appendix 7A provides details on how we calculated the numbers in this table.

Returns to education are generally higher in the informal sector for those with the lower education levels and higher in the formal sector for those with higher education levels. Conform to earlier findings (Psacharopoulos 1985), returns to education for females are higher than for males. For both males and females, an intermediate (INTER) education has higher returns if the individual participates in the informal sector. For "medio" education, males obtain higher returns in the informal sector while females' returns are higher in the formal sector. For all other education levels the returns in the formal sector exceed those in the informal, where the difference is the largest for females. Teachers training (normal) only has positive returns in the formal sector. Investments in vocational and university training have the most promising prospects on the labour market. The highest returns in the formal sector are from these, higher, education levels. This is contrary to earlier findings for developing countries where usually primary education yielded the highest returns. Not completing education has more than average negative

Table 7.1 Private returns to education (standard errors between parentheses, percent per year)

years of educ		males		females	
at completion		formal	informal	formal	informal
inter	8	1.89	2.61	5.22	7.64
		(1.23)	(1.37)	(1.74)	(1.68)
medio	12	3.37	3.92	5.61	3.52
		(0.44)	(0.50)	(0.63)	(0.64)
midtech	13	5.44	4.05	10.37	5.11
		(0.74)	(1.04)	(0.77)	(1.01)
hightech	15	5.95	2.35	9.96	4.25
		(0.79)	(1.01)	(0.90)	(1.39)
normal	17	1.80	-3.15	8.95	0.44
		(0.48)	(0.75)	(0.46)	(0.75)
university	20	5.73	2.39	9.09	3.63
		(0.24)	(0.34)	(0.35)	(0.50)
missed years		7.24	2.43	8.57	-0.31
		(0.86)	(1.14)	(1.13)	(1.33)

returns on the formal labour market for males, supporting the hypothesis that diplomas are used to screen applicants. For males in the informal and females in the formal sector, the negative returns are of the same order of magnitude as the average rates of return for other school types, supporting the human capital theory view of education. For females the negative returns are insignificant in the informal sector.

7.3 Labour Market Segmentation

Chapter 5 is the most comprehensive as far as labour market segmentation is concerned. We allowed the sector participation to depend on wage offers, sector specific non-monetary returns and rationing. Segmentation was modelled as a probability of being rationed for "good" formal sector jobs. For standard individuals (cf table 5.4) the predicted probability of rationing ranged from about 20 to 40 percent. Generally speaking, the probability of rationing decreases with additional education. Ethnic minorities and individuals living in areas with a higher unemployment rate face a higher probability of rationing.

The conventional test for labour market segmentation is based on predicted wage differences between sector. A high wage differential is associated with a low degree of competitiveness, and thus labour market segmentation. The test is derived under the assumption that individuals select their preferred sector on the basis of wage offers only. Does the conventional test lead to the same conclusions as chapter 5? To investigate this we have produced table 7.2. This table presents the predicted difference in the log wage offer in the formal and informal sector for standard individuals. A high value is associated with a relatively high formal sector wage offer and a high degree of labour market segmentation according to the conventional test. We used both models of chapter

4. The third column presents the predicted probability of rationing using the results of chapter 5. A high value is associated with a high level of labour market segmentation. The choices for the standard individuals correspond with table 5.4.

Table 7.2 Comparison of labour market segmentation tests (log formal/informal sector wage offer ratio, predicted probability of rationing, standard errors between parentheses, base: age=30,educ=inter, missed years=0, ethnic=0, econ active and unempl at mean value)

	males			females		
	probit	logit	prob rationing	probit	logit	prob rationing
base	0.72	1.24	0.33	-0.55	-1.45	0.34
	(0.08)	(0.09)	(0.02)	(0.15)	(0.16)	(0.04)
hightech educ	0.90	1.67	0.27	-0.29	-1.70	0.18
	(0.18)	(0.19)	(0.04)	(0.27)	(0.30)	(0.04)
university educ	0.92	1.86	0.20	0.08	-1.44	0.21
	(0.09)	(0.12)	(0.02)	(0.13)	(0.19)	(0.03)
ethnic=1	0.73	1.20	0.37	-0.58	-1.34	0.44
	(0.08)	(0.09)	(0.02)	(0.17)	(0.17)	(0.05)

The results in table 7.2 do not show a one to one relationship between the conventional test of labour market segmentation and the test presented in chapter 5. For both males and females, higher education increases the predicted formal sector wage relative to the informal. However, the probability of rationing decreases. In this example the two tests lead to opposite conclusions about the effect of education on labour market segmentation. One possible explanation is that education affects the non-monetary returns (this is not included in chapter 5). Lower educated might value non-monetary returns such as health insurance higher than the wealthier, well educated. Another explanation could be that there is a segmented formal labour market for high and low skilled labour. The argument would go as follows: assume that there are no returns to education in the informal sector and that in the formal sector, the wage for high skilled labour is higher than for low skilled labour. If high skilled labour is relatively scarce and unskilled labour is abundant, the informal sector wage will be determined by the equilibrium conditions for the unskilled workers. Provided that on-the-job search is relatively efficient, the resulting skilled workers will search from the informal sector. In the resulting equilibrium, the wage differential will be higher for skilled workers while at the same time the probability of entering in the formal sector is higher for skilled workers. The results in table 7.2 also indicate that belonging to an ethnic minority has no clear effect on the wage differential. The probability of rationing however, increases for both male and females. Discrimination seems to be the most plausible explanation here. In our opinion, the test results of chapter 5, based on rationing, are intuitively more appealing then those based on the wage differential. Additional research in what affects non-monetary returns seems appropriate. To this end it would be very useful if the survey would collect a monetary equivalent of some of most common forms of non-monetary compensations.

7.4 Sector Preferences

Chapter 5 models non-monetary returns associated with participation in a specific sector. We found that, on average, non-monetary returns are higher in the formal sector for males while the opposite holds for females. For the average male, a formal sector wage offer of 0.46 times (std error 0.04) that of the informal wage is required to make the (deterministic part) wage offer including non-monetary returns equate for both sectors. For females, this number is 1.18 (std err: 0.11). To get an idea of the factors driving this sector specific preference we have produced table 7A.1 (in Appendix). This table shows the effect of the taste shifters on the relative preference for the formal sector (δ_1-δ_2). We find hardly any significant effect of the family composition variables on the sector preference. Only for females, the presence of other prime age individuals in the household, increases the relative preference for the formal sector. Other income (net dissaving, to be precise) has a positive effect on the relative formal sector preference for both sexes. Both results seem to suggest that additional income sources in the household reduce the preference of working in the informal sector more than the preference for working in the formal sector. Marriage, on the other hand, has a stronger negative effect on the formal sector preference for females (see table 5.3).

7.5 Intra Household Effects

In chapter 6 we introduced a model to investigate the magnitude of intra-household effects in labour supply. We modelled the labour supply behaviour of two-adult families. We allowed for non-monetary returns to affect the sector participation decision. Simulation exercises show that the males' participation rate is not sensitive to his own wage offer (discussion based on tables 6.7 and 6.8). About 96 percent of the males work. A 10 percent drop in males' wage offers has a positive effect on hours worked of high earning males and the opposite effect for males with low wage offer. The average effect on hours worked is small. Females' hours worked increase by 3 percent. A 10 percent drop in males' wage offers has a stronger effect on female labour supply than a similar drop of the females' wage offers. We have simulated the effect of an overall reduction in formal sector wages by 10 percent. For males the result is mostly a switch between sectors and not a decrease in the participation rate. For females, average hours worked increases by 2 percent and the participation rate by 0.4 percent. The total effect on employment is positive, the fall in formal sector employment would be more than offset by a rise in informal sector participation. The size of the informal sector is expected to increase by 5.9 percent. This shows that intra-household effects are significant. In chapter 5, where we did not include intra-household effects, a reduction in the wage offer for females was predicted to lead to a lower female participation rate.

7.6 Final Remarks

The view that employment in the informal sector is inferior to formal sector employment is only partly true in the Bolivian urban labour market. For males, we do find some

support for this hypothesis. Most males prefer a formal sector job and expected wages are higher than in the informal sector. Barriers to entry seem to be especially relevant for the lower educated and for individuals belonging to ethnic minorities. For females, however, we find that on average, participation in the informal sector is preferred to a formal sector job and that expected hourly earnings are higher in the informal sector. Barriers to entry in the formal sector are similar to those for males, but less relevant since the informal sector is usually preferred. The fact that most working females participate in the informal sector while the opposite holds for males seems to be supply driven. There is no evidence for sex-discrimination in the formal sector. The description of the informal sector as a buffer sector, inferior to formal sector employment, only seems to be appropriate for males. This is in line with the result of Magnac (1991) who found no evidence of labour market segmentation for married Columbian females.

Appendix 7A

Returns to Education. The numbers appearing in table 7.1 are based on the estimates for the wage equations in chapters 4 and 5. Chapter 4 contains two estimates, chapter 5 one. For the formal sector in chapter 5 we have used the results for the "good" formal sector. The estimates are different because we corrected for selection bias in a different way, because we included a distinction between good and bad formal sector jobs only in chapter 5 and because the sector definition has changed from chapter 4 to 5. However, the estimates are correlated as they are based on the same dataset. The specification of the wage equation is the same in all three models. Table 7.1 is constructed as follows: first, we estimated the coefficients for the wage equations by asymptotic least squares (ALS). Second, we divided the coefficients on the education dummies in the wage equation by the corresponding years of education. Because "basic" is the excluded category in the wage equations we divided by number of years of education belonging to each school type minus 5 (the number of years needed to complete basic education). The years of education belonging to each type of training are in appendix I. The figures in table 7.1 are thus the average rates of return for one year of education after "basic" for different educational patterns.

Let $\hat{\beta}=(\hat{\beta}_1,\hat{\beta}_2,\hat{\beta}_3)$ be estimates for the coefficients in the wage equations in chapters 4 and 5. The ALS estimate, $\hat{\beta}_{ALS}$, is obtained by

$$\hat{\beta}_{ALS} = \underset{\beta}{argmin}\ f(\beta)' W f(\beta) \quad with\ f(\beta) = \begin{bmatrix} \beta-\hat{\beta}_1 \\ \beta-\hat{\beta}_2 \\ \beta-\hat{\beta}_3 \end{bmatrix} \tag{7A.1}$$

The optimal weighting matrix W, is the inverse of the covariance matrix of $\hat{\beta}$. In each of the models β_i is a subvector of the full parameter vector. Denote the latter by $\theta_i=(\theta_{1i},\theta_{2i})$ with $\theta_{2i}=\beta_i$. The asymptotic distribution of $\hat{\beta}_i$ satisfies

$$\sqrt{N}(\hat{\beta}_i-\beta) \cong B_{12}\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\frac{\partial \log p_{ij}}{\partial \theta_{1i}}\Big|_{\theta_i} + B_{22}\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\frac{\partial \log p_{ij}}{\partial \theta_{2i}}\Big|_{\theta_i} \tag{7A.2}$$

with

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{bmatrix}^{-1} \quad with\ A_{kl}=\frac{1}{N}\sum_{j=1}^{N}\left[\left(\frac{\partial \log p_{ij}}{\partial \theta_{ki}}\right)\left(\frac{\partial \log p_{ij}}{\partial \theta_{li}}\right)'\right] \tag{7A.3}$$

and p_{ij} the likelihood contribution for observation j in model i. By standard argument the expression for the covariance matrix of $\hat{\beta}$ is obtained using (7A.2) for i=1,2,3. The covariance matrix of β_{ALS} is $(\Gamma'W\Gamma)^{-1}$ with $\Gamma=\partial f(\beta)/\partial\beta$.

Table 7A.1 Determinants of formal sector vs informal sector preferences (δ_1-δ_2; see formula (5.5), standard errors in parentheses)

	males	females
cnst	2.562	1.078
	(0.410)	(0.483)
age	-0.059	-0.021
	(0.019)	(0.024)
age squared /100	0.020	-0.026
	(0.023)	(0.029)
young	0.017	-0.015
	(0.019)	(0.021)
prime	0.021	0.072
	(0.020)	(0.021)
old	-0.149	0.066
	(0.088)	(0.105)
married	-0.866	-1.188
	(0.239)	(0.232)
net pc dissavings /1000	1.843	3.458
	(0.780)	(0.809)
pred error	-0.974	-3.079
	(0.762)	(0.816)
married*age	0.022	0.018
	(0.006)	(0.006)

Appendix I. List of Variables

Individual level variables

married	dummy =1 if individual is married
ethnic	dummy =1 if the individual regularly speaks other than spanish
active search	if the individual reported to be searching for work in the last week.
discouraged	if the individual reported not to be looking for work because "there is no work"
hourly earnings	calculated as total earnings divided through the number of hours worked (sometimes also referred to as wage)

Education variables

variable dummies	description	number of years of education at completion
Basic	includes no education and basic education	5
inter	intermediate education	8
medio		12
midtech	vocational training, includes industry, commercial and agriculture training.	13
hightech		15
normal	primary school teacher training	17
university	university both private and public	20
other	non-categories, for males probably largely military training	
missed years	if individual did not complete training: minus number of years before completion	

Household level variables

net pc dissavings	total monthly household consumption minus total household labour income divided by family size
other income	total household non-labour income per month
young	number household members of less than 15 years old
prime	number of household members between 15 and 65
old	number of household members greater than 65 years old

Regional variables

unemployment	share of labour force unemployed per urban center
econ act	number of working or searching individuals in the sample per urban center

References

BLUNDELL, R. (1990), "Evaluating Structural Microeconometric Models of Labour Supply", Working Paper, University College London.

BLUNDELL, R. AND I. WALKER (1986), "A Life Cycle Consistent Empirical Model of Labour Supply using Cross Section Data", Review of Economic Studies 53, pp. 539-558.

BOERSCH-SUPAN, A. AND V. HAJIVASSILIOU (1993), "Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models," Journal of Econometrics, vol. 58, no 3, pp. 347-368.

BROCK,W AND D. EVANS (1986), The Economics of Small Business: their role and regulation in the US economy, Holmes & Meier, New York, London.

CONTRERAS, M. (1990), "Debt, Taxes, and War: The Political Economy of Bolivia, c. 1920-1935", Journal of Latin American Studies, vol. 22, no 2, pp. 265-287.

DEVINE, T.J. AND N.M. KIEFER (1991), Empirical Labor Economics: the Search Approach, New York: Oxford University Press.

DEWIT, G. (1991), Determinants of Self-Employment, PhD thesis, University of Amsterdam.

DICKENS, W.T. AND K. LANG (1985), "A Test of Dual Labor Market Theory", The American Economic Review, vol 75, no 4 pp. 792-804.

DICKENS, W.T. AND K. LANG (1992), "Labor Market Segmentation Theory: Reconsidering the Evidence", NBER working paper nr 4087, Cambridge MA, USA.

DICKENS, W.T. AND S.J. LUNDBERG (1993), "Hours restrictions and labor supply", International Economic Review, vol. 34, pp. 169-192.

DOMENCICH, T., AND D. MC FADDEN (1975), Urban Travel Demand: A behavioral Analysis, Amsterdam: North Holland.

FIELDS, G.S. (1975), "Rural-Urban Migration, urban Unemployment and Underdevelopment, and job search activity in LDCs'", Journal of Development Economics, vol. 2, pp. 165-187.

FIELDS, G.S. (1989), "On the Job Search in a Labor Market Model: Ex ante Choices and Ex Post Outcomes'", Journal of Development Economics, vol. 30, pp. 159-178.

FIELDS, G.S. (1990), "Labour Market Modelling and the Urban Informal Sector: Theory and Evidence", in D. Turnham, B Salomé and A. Schwartz (eds), The informal Sector Revisited, OECD Development Centre: Paris.

GAAG, J. VAN DER AND W. VIJVERBERG (1988), A Switching Regression Model for Wage Determinants in the Public and Private Sectors of a Developing Country, The Review of Economics and Statistics 70, no. 2, pp. 244-252.

GINDLING, T.H. (1991), "Labor Market Segmentation and the Determination of Wages in the Public, Private-Formal, and Informal Sectors in San José, Costa Rica", Economic Development and Cultural Change, vol 39, no 3, pp. 585-606.

GOURIEROUX, C. AND A. MONFORT (1993), "Simulation based Inference: A Survey with Special Reference to Panel Data Models", Journal of Econometrics, vol. 59, pp. 5-34.

GOURIEROUX, C. AND A. MONFORT (1989), Statistiques et Modeles Econometriques, Paris: Economica.

HARRIS, J.R. AND M.P TODARO (1970), "Migration, Unemployment and Development: a two Sector Analysis", American Economic Review, vol. 60, pp. 126-142.

HART, K. (1985), The Informal Economy, Cambridge Anthropology 10, no 2, 54-58.

HARTOG, J. AND H. OOSTERBEEK (1993), Public and Private Sector Wages in the Netherlands, European Economic Review 37, pp. 97-114.

HAUSMAN, J. AND P. RUUD (1984), "Family labor supply with taxes", American Economic Review, vol. 74, pp. 242-248.

HAUSMAN, J.A. (1978), Specification Tests in Econometrics, Econometrica 46, pp. 1251-1272.

HECKMAN, J. (1974), "Shadow Prices, Market Wages and Labor Supply", Econometrica, vol. 47, pp. 153-161.

HECKMAN, J. (1978), "Dummy Endogenous Variables in a Simultaneous Equation System", Econometrica, vol. 46, pp. 931-960.

HECKMAN, J. (1990), Varieties of Selection Bias, American Economic Review 80, no 2, 313-318.

HECKMAN, J. AND B.E. HONORÉ (1990), The Empirical Content of the Roy Model, Econometrica 58, no 5, pp. 1121-1149.

HECKMAN, J. AND V.J. HOTZ (1986), The Sources of Inequality for Males in Panama's Labor Market, The Journal of Human Resources 21, no 4, pp. 507-542.

HECKMAN, J. AND V.J. HOTZ (1986), "An Investigation of the Labor Market Earnings of Panamanian Males: Evaluating the Sources of Inequality", Journal of Human Resources, 21(4), Fall 1986, pp. 507-42.

HORTON, S. (1994 forthcoming), "Labor Markets in an Era of Adjustment: Bolivia", in S Horton, R. Kanbur and D. Mazumdar (eds), Labor Markets in an Era of Adjustment, Washington D.C.: World Bank EDI.

IDSON, T.L. AND D.J. FEASTER (1990), "A Selectivity Model of Employer Size Wage Differentials", Journal of Labor Economics, vol. 8, no 1, pp.99-122.

INTERNATIONAL LABOUR OFFICE (1972), Employment, Incomes and Equality: A Strategy for Increasing Productive Employment in Kenya, Geneva.

INTERNATIONAL LABOUR OFFICE (1991), Yearbook of Labour Statistics, 50 ed., Geneva.

JEMIO, L.C. (1993), Micro- and Macroeconomic Adjustment in Bolivia (1970-89), PhD thesis, Institute for Social Studies, The Hague, the Netherlands

KAPTEYN, A., P. KOOREMAN AND A. VAN SOEST (1990), "Quantity rationing and concavity in a flexible household labor supply Model", The Review of Economics and Statistics, vol. 70, pp. 55-62.

LAROQUE, G. AND B. SALANIE (1989), "Estimation of multi-market fix-price models: An application of Pseudo Maximum Likelihood methods", Econometrica, vol. 57, no. 4, pp. 831-860.

LEE, L.F. (1982), Some Approaches to the Correction of Selectivity Bias, Review of Economic Studies 49, pp. 372.

LERMAN, S. AND C. MANSKI (1981), On the Use of Simulated Frequencies to Approximate Choice Probabilities in : C. Manski and D. McFadden (eds), Structural Analysis of Discrete Data with Econometric Applications, Cambridge, Mass, MIT press.

LINDERT, P. VAN (1990), Huisvestingsstrategiën van lage-inkomensgroepen in La Paz, Nederlands Geografische Studies 136, Utrecht.

LUBELL, H. (1990), The Informal Sector in the 1980s and 1990s, Development Centre O.E.C.D., Paris.

MACURDY, T., D. GREEN AND H. PAARSCH (1990), "Assessing empirical approaches for analyzing taxes and labor supply", Journal of Human Resources, Vol. 25, pp. 415-490.

MADDALA, G.S. (1983), Limited Dependent and Qualitative Variables in Econometrics, 1 ed., Cambridge University Press, Cambridge.

MAGNAC, TH. (1991), "Segmented or Competitive Labor Markets", Econometrica, vol. 59, pp.165-187.

MCFADDEN, D. (1974), "The Measurement of Urban travel Demand", Journal of Public Economics, vol. 3, pp.303-328.

MEZZERA, J. (1985), Segmentación del Mercado Liberal, PREALC, Santiago, Chile.

MINCER, J. (1974), Schooling, Experience and Earnings, National Bureau of Economic Research, New York.

MINCER, J. (1976), "Unemployment Effects of Minimum Wages", Journal of Political Economy, vol. 84, no 4, pp.s87-s104.

MORALES, W.Q. (1992), Bolivia: Land of Struggle, Westview Press, Boulder, U.S.A.

MORALES, J.A. AND J. SACHS (1988), "Bolivia's Economic Crisis", NBER working paper nr 2620, Cambridge MA, USA.

NEWEY, K.N. (1984), "A Method of Moment Interpretation of Sequential Estimators", Economic Letters 14, pp. 206.

NEWMAN J.L. AND P.J. GERTLER (1992), "Family productivity, Labor Supply and Welfare in a low Income Country", working paper, World Bank (PHRWH), Washington D.C.

NEWMAN, J, S. JORGENSEN AND M. PRADHAN (1991), "How Did Workers Benefit from Bolivia's Emergency Social Fund?", The World Bank Economic Review vol 5, no 2, pp. 367-393.

PRADHAN, M.P. (1993), "Sector Participation in Labour Supply Models: preferences of Rationing", CentER discussion paper No 9366, Tilburg University.

PRADHAN, M.P. AND A.H.O. VAN SOEST. (1992), "Formal and Informal Sector Employment in Urban Areas of Bolivia", CentER discussion paper No 9311, Tilburg University, and Labour Economics, forthcoming.

_____ (1994), "Household Labour Supply in Urban Areas of a Developing Country", CentER discussion paper No 9429, Tilburg University.

PRICE WATERHOUSE (1988), Doing Business in Bolivia, La Paz , Bolivia.

PSACHAROPOULOS, G. (1985), "Returns to Education: A Further International Update and Implications", The Journal of Human Resources, vol. 20, no 4, pp.583-604.

RANSOM, M.R. (1987), "An Empirical Model of Discrete and Continuous Choice in Family Labor Supply", The Review of Economics and Statistics, vol. 69, pp. 465-472.

RODRIGUEZ, R.S. AND A.R. ARENAS DE SANDOVAL (1992), Legislación del Trabajo, La Paz , Bolivia.

ROSENZWEIG, M.R. (1988), "Labor Markets in Low-Income Countries", in H. Chenery ,T.N. Srinivasan (eds), Handbook of Development Economics, Vol 1, North Holland: Amsterdam.

ROY, A.D. (1951), "Some Thoughts on the Distributions of Earnings", Oxford Economic Papers (New Series), vol. 3, pp. 135-146.

SACHS, J. (1987), "The Bolivian Hyperinflation and Stabilization", The American Economic Review, vol. 77, no 2, pp. 279-283.

SAINZ, R.C. AND S.E. DE PABÓN (1988), Los Trabajadores por Cuenta Propia en La Paz, CEDLA, La Paz , Bolivia.

SCOTT, D.W. (1992), Multivariate Density Estimation, Wiley, New York.

SMITH, R.J. AND R. BLUNDELL (1986), An exogeneity test for a simultaneous equation Tobit model with an application to labor supply, Econometrica 54, pp. 679-686.

STERN, S. (1992), "A Method for Smoothing Simulated Moments of Discrete Probabilities in Multinominal Probit Models", Econometrica, vol. 60, no 4, pp. 943-952.

STRASSMANN, W.P. (1987), "Home-Based Enterprises in cities in Developing Countries", Economic Development and Cultural Change, vol. 36, pp. 121-144.

SUMMERS,R. AND A.M. HESTON (1988), "A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries, 1950-1985", Review of Income and Wealth, Vol 34, No 1, pp. 1-25.

TANNEN, M.B. (1991), "Labor Markets in Northeast Brazil: Does the Dual Market Model Apply?", Economic Development and Cultural Change, vol 39(3), pp. 567-83.

THOMAS, J.J. (1992), Informal Economic Activity, LSE Handbooks in Economics, Hertfordshire: Harvester Wheatsheaf.

TODARO, M.P. (1989), Economic Development in the Third World, New York: Longman.

U.D.A.P.E. (1991), Estadísticas Económicas de Bolivia, nr 2, La Paz, Bolivia.

U.D.A.P.E. (1992), Estadísticas Económicas de Bolivia, nr 3, La Paz, Bolivia.

UNIDAD DE ANALISIS DE POLITICAS ECONOMICAS (1991), Estadísticas Economicas de Bolivia, La Paz.

VAN PRAAG, B.M.S. AND J.P. HOP (1987), "Estimation of Continuous Models on the Basis of Set-valued Observations", mimeo, Erasmus University Rotterdam.

VAN SOEST, A., P. KOOREMAN AND A. KAPTEYN (1993), "Coherency and regularity of demand systems with equality and inequality constraints", Journal of Econometrics, vol. 57, pp. 161-188.

VELASCO, A.P., R.C. SAINZ, S.E. PABON, Y H.L. CORDOVA (1989), Informalidad e Ilegalidad: Una Falsa Indentidad, ed., CEDLA, La Paz.

VUONG, Q.H. (1989), "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses", Econometrica, vol. 57, pp. 307-333.

WELLINGS, P AND M. SUTCLIFFE (1984), "Developing the Urban Informal Sector of Developing Economies", Development and Change, vol. 15, pp. 517-550.

WORLD BANK (1986), Updating Economic Memorandum on Bolivia, report nr 6455-BO Washington DC, USA.

WORLD BANK (1988), Bolivia Updating Economic Memorandum, report nr 7278-BO Washington DC, USA.

WORLD BANK (1989), Bolivia Updating Economic Memorandum, report nr 7645-BO Washington DC, USA.

WORLD BANK (1990a), Bolivia Updating Economic Memorandum, report nr 8623-BO Washington DC, USA.

WORLD BANK (1990b), Bolivia Poverty Report, report nr 8643-BO Washington DC, USA.

WORLD BANK (1991), Bolivia: From Stabilization to Sustained Growth , report nr 9763-BO Washington DC, USA.

WORLD BANK (1992a), Bolivia Updating Economic Memorandum, report nr 11123-BO Washington DC, USA.

Samenvatting

In dit proefschrift wordt de rol van de informele sector in stedelijke gebieden van ontwikkelingslanden, en met name Bolivia, onderzocht. De informele sector wordt vaak gezien als een buffersector. Volgens deze visie wordt werk in de formele sector geprefereerd boven werk in de informele sector. Echter, niet iedereen kan tot de formele sector toetreden. Als een gevolg van restricties op de arbeidsmarkt, zoals bijvoorbeeld een wettelijk minimum loon, is werk in de formele sector gerantsoeneerd. De informele sector, daarentegen, is voor iedereen toegankelijk. Werken in de informele sector wordt gezien als tweede keus, met lagere opbrengsten dan in de formele sector. De laatste jaren zijn er twijfels gerezen ten aanzien van deze hypothese. Uit empirische studies blijkt dat een groot gedeelte van de informele sector werkers er de voorkeur aan geeft in deze sector te blijven. In deze studie proberen we aan te geven wat de rol van de informele sector is in Bolivia. Het onderzoek is gebaseerd op gegevens van een huishoudensenquête, uitgevoerd in 1989 onder 7246 huishoudens in stedelijke gebieden. De analyse wordt uitgevoerd met behulp van statische arbeidsaanbodmodellen. We vinden dat de bufferfunctie een goede beschrijving is voor mannen. Echter, voor vrouwen vinden we dat de informele sector vaak geprefereerd wordt ten opzichte van de formele sector.

De centrale vraagstelling is in hoeverre huishoudens schokken in de formele sector zullen opvangen door middel van een reallocatie van hun arbeidsaanbod. Structurele aanpassingsprogramma's hebben vaak directe negatieve consequenties voor beloningen in de formele sector. De daaruit volgende inkomensdaling van het huishouden kan worden opgevangen door meer te gaan werken en/of over te stappen naar de informele sector. Ook kunnen er binnen het huishouden aanpassing plaatsvinden. De vrouw kan bijvoorbeeld besluiten te gaan werken als de man wordt getroffen door een inkomensdaling. De uiteindelijke gevolgen voor huishoudconsumptie zullen waarschijnlijk lager uitvallen dan het initieel gederfde inkomen. In dit onderzoek zullen we trachten deze effecten te quantificeren door arbeidsaanbodelasticiteiten te schatten op basis van cross sectie data. De analyse is partieel in de zin dat we alleen rekening houden met arbeidsaanbodeffecten. Veranderingen in de vraag naar arbeid worden buiten beschouwing gelaten.

Hoofdstuk 1 bevat een inleiding en een korte samenvatting van wat zal volgen.

Het empirische gedeelte van dit proefschrift heeft betrekking op Bolivia. Hoofdstuk 2 bevat achtergrondinformatie betreffende de economische geschiedenis, het belastingstelsel, de regelgeving met betrekking tot de arbeidsmarkt en het datamateriaal dat zal worden gebruikt in de empirische hoofdstukken. Na een periode van hyperinflatie in het begin van jaren tachtig werd de economie in 1985 gestabiliseerd. Het aanpassingsprogramma was rigoureus en had grote consequenties voor de arbeidsmarkt. De informele sector groeide met 14 procent. In 1986 was 60 procent van de werkenden in de informele sector. De daarop volgende jaren werden gekenmerkt door een geringe

groei en een relatief stabiele economie. Het leeuwedeel van de belastingen wordt geheven door middel van een toegevoegde waarde belasting. Inkomstenbelastingen spelen een geringe rol, slechts 3 procent van de totale overheidsinkomsten wordt via deze belastingen verkregen. We zullen daarom geen rekening houden met inkomstenbelasting bij het modelleren van de arbeidsaanbodbeslissing in de empirische hoofdstukken.

Hoofdstuk 3 bevat een overzicht van de literatuur betreffende arbeidsmarkten in ontwikkelingslanden. Het eerste gedeelte geeft een overzicht van de theoretische modellen. De eerste modellen komen voort uit de migratieliteratuur. In het begin van de jaren zeventig werden deze modellen uitgebreid met een informele sector. In deze modellen werd de informele sector gezien als een springplank naar de formele sector: vanuit de informele sector kon worden gezocht naar een baan in de formele sector. Empirische studies, die in het tweede gedeelte van dit hoofdstuk worden behandeld, geven echter aan dat deze hypothese niet altijd opgaat. Een substantieel gedeelte van de informele sector werkers is niet op zoek naar een formele sector baan. Studies waarbij lonen tussen de twee sectoren worden vergeleken geven geen eenduidige conclusie. In het algemeen wordt gevonden dat de beloningsstructuur in de formele en informele sector verschillen. De conclusies betreffende arbeidsmarktsegmentatie, ofwel rantsoenering van formele sector banen, verschillen per studie. Tenslotte worden in dit hoofdstuk een aantal suggesties gedaan voor verbetering van de theoretische en empirische modellen. Het belangrijkste euvel van de theoretische modellen is dat de informele sector homogeen wordt verondersteld. We geven een suggestie voor een alternatief model met een ongelijke inkomensverdeling in de informele sector. Verbeteringen van de empirische modellen volgen in de latere hoofdstukken.

In hoofdstuk 4 analyseren we participatie en beloningen in de formele en informele sector in Bolivia in 1989. Een bekend probleem bij het schatten van loonfuncties is dat simpelweg kleinste kwadraten toepassen niet consistent is indien er sprake is van niet geobserveerde karakteristieken die zowel de lonen als de sectorparticipatie beïnvloeden. Voor modellen waarbij de keuze slechts uit twee alternatieven (werken/niet werken) bestaat, is een uitgebreide literatuur voorhanden. Echter, in ons geval is er sprake van drie keuzes: Werken in de formele sector, in de informele sector en niet werken. We experimenteren met twee modellen om de participatie te modelleren. In het eerste model veronderstellen we dat de informele sector een buffersector is, tussen niet werken en werken in de formele sector. In het tweede model wordt geen a priori ordening tussen sectoren verondersteld. Op basis van specificatietoetsen hebben we een lichte voorkeur voor het eerste model voor mannen en het tweede voor vrouwen. Verder vinden we dat de keuze van het participatiemodel een grote invloed heeft op de schattingen voor de loonfuncties. In het algemeen vinden we dat de voorspelde lonen hoger zijn in de formele sector dan in de informele sector voor mannen terwijl voor de vrouwen de voorspelde lonen hoger zijn in de informele sector. Dit laatste is verrassend. In andere opzichten zijn de resultaten conform de verwachtingen: opbrengsten van scholing zijn hoger in de formele sector dan in de informele sector en zijn hoger voor vrouwen dan

voor mannen, de informele sector is gevoeliger voor lokale arbeidsmarkt omstandigheden en etnische minderheden worden slechter betaald.

Een toets voor arbeidsmarktsegmentatie gaat in het algemeen als volgt: op basis van geobserveerde en niet geobserveerde karakteristieken worden de lonen in de formele en informele sector voorspeld. Als het voorspelde loon in de formele sector consequent hoger is dan in de informele sector wordt dit geïnterpreteerd als een teken van arbeidsmarktsegmentatie. Echter, het is ook mogelijk dat loonverschillen voortkomen uit sectorspecifieke preferenties van individuen. Het kan bijvoorbeeld zo zijn dat de vrijheid die de informele sector biedt, wordt gewaardeerd en meegenomen in de sectorparticipatie keuze. In dat geval zou het loon in de informele sector lager kunnen zijn, zonder dat er sprake is van rantsoenering in de formele sector. Het is dus belangrijk de effecten van sector specifieke preferenties te kunnen onderscheiden van rantsoenering bij het toetsen van arbeidsmarktsegmentatie. Hoofdstuk 5 presenteert een model waarbij de effecten van preferenties en rantsoenering in de participatie beslissing apart worden gemodelleerd. Dit gebeurt door gebruikt te maken van zoekinformatie in de dataset. De veronderstelling is dat zoeken naar werk kan worden gezien als een indicatie van rantsoenering. Indien iemand op zoek is naar werk geeft dat aan dat hij of zij niet tevreden is met de huidige situatie en verwacht een betere baan te kunnen vinden. Het is echter ook mogelijk dat iemand het zoeken heeft opgegeven omdat de kans op beter werk te klein was. Deze mensen zijn ook gerantsoeneerd en wij hebben hen daarom identiek behandeld als actieve zoekers. Nu we rantsoenering hebben gemodelleerd kunnen we het overige gedeelte van de voorspelde loonverschillen toeschrijven aan preferenties. We vinden dat, gemiddeld gezien, mannen de formele sector prefereren en vrouwen de informele. De kans op rantsoenering voor formele sector banen ligt voor beide sexen rondom 30 procent. De relatief hoge participatie van vrouwen in de formele sector is dus toe te schrijven aan verschillen in preferenties, en niet het gevolg van discriminatie van vrouwen voor formele sector banen.

De arbeidsaanbodbeslissing van een individu kan worden beïnvloed door de arbeidsmarkt positie en inkomen van overige leden in het huishouden. Bijvoorbeeld, indien de man in een familie een salarisverlaging krijgt kan de vrouw besluiten te gaan werken om het gederfde inkomen te compenseren. In hoofdstuk 6 analyseren de arbeidsaanbodbeslissingen van huishoudens met één volwassen man en één volwassen vrouw. In het model veronderstellen we dat de arbeidsaanbodbeslissingen simultaan worden genomen, dat wil zeggen: de man en de vrouw beslissen gezamenlijk over wie gaan werken en hoeveel uren zij gaan werken. We onderscheiden weer de formele en de informele sector. Rantsoenering en sectorspecifieke preferenties worden gezamenlijk gemodelleerd en aangeduid met "niet-monetaire beloningen" van werk. We vinden substantiële intra-huishoudeffecten. Lagere lonen van de man worden gecompenseerd door extra werk van de vrouw. Een gesimuleerde loonsdaling in de formele sector van 10 procent geeft een positief effect op de algehele participatie. De participatie van mannen neemt iets af, vrouwen gaan echter meer werken. De informele sector groeit met 6 procent. Dit moet worden gezien als een overschatting van het werkelijke effect. Indien

rekening zou worden gehouden met algemeen evenwicht effecten, en de daling van de vraag naar goederen uit de informele sector zou worden meegemodelleerd, zou de toename lager uitvallen.

Hoofdstuk 7 bevat conclusies.